Analysis of the Pumilio Repression Mechanism and its Impact on the Transcriptome

A Dissertation
SUBMITTED TO THE FACULTY OF
THE UNIVERSITY OF MINNESOTA
BY

Rebecca J. Haugen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor Dr. Aaron Goldstrohm
Committee members: Dr. Anja Bielinsky, Dr. Eric A. Hendrickson, Dr. David Greenstein

May 2021

Acknowledgements

I would like to begin by acknowledging my husband Travis, for his ceaseless and unwavering support on this long journey. None of this would have been possible without him. I also thank my kids, Cody, Sarah, and Grace, for sacrificing some of their “Mom time” to all-day writing sprees or late nights in the lab. To Cody, for showing a keen interest in the science and allowing me to practice explaining it to him; to Sarah, for being my angel of sanity with relentless encouragement; to Grace, for providing comedic relief, and also engaging in late-night science talks: you have all been more instrumental in this accomplishment than I could ever explain. I would also like to thank my parents for their support and encouragement.

I would like to thank my advisor Dr. Aaron Goldstrohm, for his guidance, and seemingly bottomless reservoir of patience and flexibility. I also thank my committee members Drs. Anja Bielinsky, Eric A. Hendrickson, and David Greenstein, for their encouragement, help, and feedback. I would like to especially thank Dr. Greenstein for supporting me and believing in me.

I want to also thank the members of the Goldstrohm lab, past and present for their insight, encouragement, and scientific discussion, especially Dr. Katie McKenney, and my bench-mate, Rob Connacher.

Dedication

This work is dedicated to my daughters, Sarah and Grace, and my son Cody. May you find enjoyment in working hard for something.

Abstract

The expression of genes in an organism is controlled by many methods. One of the most important means of regulation is post-transcriptional control of messenger RNA (mRNA). RNA binding proteins (RBPs) are crucial for this process. Here we describe the characterization of a functional domain within the *Drosophila* RBP Pumilio (Pum) and examine the effects of Pum on the transcriptome.

Pum proteins are a highly conserved family of RBPs that control the expression of mRNA by binding to specific motifs in the 3'-UTR of transcripts. Through associations with other proteins, Pum proteins regulate the fate of an RNA, primarily causing decay of the target. *Drosophila* Pum contains autonomous domains capable of enacting decay independently from its RNA binding domain. The most conserved domain is repression domain 3 (RD3) which we characterize here. Using conservation analysis, we define conserved regions within RD3, and show their importance for the repressive activity of RD3. We identify two specific amino acids, F1033 and F1040, which are necessary for RD3's function. Using yeast two-hybrid assays, we show that RD3 contacts several members of the CNOT deadenylase complex, the central scaffold unit Not1, and subunits Not2 and Not3. In *Drosophila* cells, we show that RD3 can interact with the CNOT complex. We further explore the endogenous targets of *Drosophila* Pum proteins by performing RNA sequencing experiments in which we deplete Pum and members of the CNOT complex. This is the first global analysis of the functional regulation of Pum targets in *Drosophila*. From this analysis we find that thousands of genes rely upon the regulatory framework provided by Pum. We demonstrate specific regulation of several targets of Pum using reporter assays and identify functional Pum binding sites within target 3'-UTRs. Finally, we show that regulation by Pum also affects the levels of expressed protein for two important targets, Pde11 and Raf. The work shown here enhances our understanding of Pum proteins for all organisms and the field of post-transcriptional control.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Name
4E-BP	eIF4E binding protein
aa	Amino acid
AD	Activation domain (Y2H assays)
Ade	Adenine, amino acid
APA	Alternative polyadenylation
ARE	A-U rich element
AzVT	Azido nucleotide (AzATP, AzGTP, AzCTP)
BD	Binding domain (Y2H assays)
cAMP	Cyclic AMP, adenosine monophosphate
CBC	Cap-binding complex
CDS	Coding sequence
cGMP	Cyclic GMP, guanine monophosphate
Click-Seq	Click chemistry-ligated adapters with sequencing
CLIP-seq	Cross-linking immunoprecipitation with sequencing
CNOT	CCR4-Not1 deadenylase complex (Carbon catabolite repressor 4; Negative on transcription)
CPE	Cytoplasmic polyadenylation element
CPEB	Cytoplasmic polyadenylation element binding protein
CPSF	Cleavage and polyadenylation specificity factor
CR	Conserved region
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPR-Cas9	System of genome editing using Cas9 enzyme
CTD	C-terminal domain (of RNA polymerase)
DDO	Double dropout media (lacking Trp and Leu)
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide-triphosphate
dsRNA	Double-stranded RNA
eEF	Eukaryotic elongation factor
eGFP	Enhanced Green fluorescent protein
EGFR	Epidermal growth factor receptor
eIF4F	Eukaryotic initiation factor 4F
EJC	Exon junction complex
eRF	Eukaryotic release factor
FBF	Fem-3 binding factor
FBS	Fetal bovine serum
FDR	False discovery rate
FFLuc	Firefly luciferase
FPKM	Fragments per kilobase million
GO	Gene ontology
His	Histidine, amino acid
HT	HaloTag
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
Indel	Insertion/deletion

IP	Immunoprecipitation
IRS	Insulin receptor substrate
KD/KO	Knockdown/knockout
kDa	KiloDalton (molecular weight)
Leu	Leucine, amino acid
lncRNA	Long noncoding RNA
MAPK	Mitogen-activated protein kinase
miRNA	Micro RNA
MS	Mass spectrometry
MWM	Molecular weight marker
NED	Nanos effector domain
NIM	Not1 interaction motif
NLuc	Nano luciferase
NMD	Nonsense-mediated decay
NMJ	Neuromuscular junction
NTC	Non-targeting control
PABP(N/C)	Poly(a) binding protein, Nuclear or Cytoplasmic
PAC-Seq	Poly(A) Click sequencing
PAP	Poly(A) polymerase
PARN	Poly(a) ribonuclease
PAS	Polyadenylation sequence
PCMa	Pumilio conserved motif (a or b)
PDE	Phosphodiesterase
PRE	Pumilio response element
PTC	Premature termination codon
PUF	Pumilio and FBF
QDO	Quadruple dropout media (lacking His, Leu, Ade, Trp)
qPCR	Quantitative PCR
RBD	RNA-binding domain
RBP	RNA binding protein
RD (1, 2, or 3)	Repression domain 1, 2, or 3
RIP	RNA immunoprecipitation
RIP-Chip	RNA immunoprecipitation with micro-array analysis
RNA	Ribonucleic acid
RNA Pol	RNA polymerase
RNAi	RNA interference
RNP	Ribonucleoprotein
RT	Reverse transcription
RT-qPCR	Reverse transcription- quantitative PCR
SEM	Standard error of the mean
tRNA	Transfer RNA
Trp	Tryptophan, amino acid
UTR	Untranslated region
WT	Wild-type
Y2H	Yeast two-hybrid

Chapter 1: Introduction

The precise regulation of gene products is crucial during an organism's development and entire life. Even small perturbations in the proper level of gene expression can result in disease states or developmental abnormalities. Several mechanisms exist to maintain the appropriate balance of proteins, such as control of gene transcription, spatial and temporal localization of translated proteins, and protein degradation. Another important regulatory step is the management of transcribed RNAs, in the form of post-transcriptional control.

Development of all animals is dependent on a system of post-transcriptional control. Maturation of a fertilized oocyte into a zygote is a transcriptionally silent process that relies on maternally deposited proteins and RNAs. Until transcription of the zygotic genome begins, oocyte cell cycle completion, cell division, and localization of cell fate determinants must all be regulated post-transcriptionally (Tadros and Lipshitz 2009).

An organism must also be able to respond to internal and external stimuli, sometimes immediately. Most biological processes that occur are carried out by proteins, yet the process of making a protein is time and energy consuming. One of the ways cells handle this dilemma is by transcribing genes compulsorily then exerting control over the product RNA. The RNA can then be translated into protein, stored to mount a quicker response at a later time, or transported to a location where the final protein will ultimately function. Post transcriptional control allows for exquisite fine tuning of cellular responses.

1.1. Making the mRNA

The first step in the life cycle of all RNA is transcription. DNA transcription in eukaryotes is carried out by one of three RNA polymerases (RNA pol), each one responsible for transcribing a certain subset of genes. RNA transcripts that are destined to become proteins are transcribed by RNA pol II. An extended C-terminal domain (CTD) on RNA pol II serves as a scaffold for factors unique to processing pre-mRNAs such as capping and splicing complexes. This spatial arrangement allows for RNA processing

events to occur co-transcriptionally (Young 1991, Willis 1993, Cho, Takagi et al. 1997, McCracken, Fong et al. 1997, Bentley 2002, Dieci, Fiorino et al. 2007).

Capping of the pre-mRNA occurs immediately after transcription elongation begins. The cap consists of a single guanine nucleotide appended to the end of the transcript with an unusual 5' to 5' triphosphate linkage. A capping complex is recruited to the CTD of RNA pol II after the phosphorylation of the RNA pol II CTD (Gonatopoulos-Pournatzis and Cowling 2014). Capping occurs in 3 stages: removal of the gamma phosphate group from the 5' terminal nucleotide to yield a 5' di-phosphate; a guanylyl transferase then hydrolyzes GTP to transfer a GMP group to the transcript, finally a methyl group is added to the N7 position using s-adenosylmethionine (Banerjee 1980). In yeast, three separate enzymes perform these processes, but in metazoans, the phosphatase and guanylyl transferase activity are combined into one enzyme- RNGTT in humans (Pillutla, Shimamoto et al. 1998, Yamada-Okabe, Doi et al. 1998) and mRNA-cap in *Drosophila*.

Capping of RNA serves several important functions, the first of which is to protect the nascent RNA from enzymatic cleavage. The unique triphosphate linkage between the cap and the 5' end of the transcript requires a specialized enzyme for cleavage. Uncapped transcripts are rapidly degraded; even viruses have evolved ways of stabilizing their own mRNAs with caps or stealing host caps in order to evade destruction of their messages. 5' caps in eukaryotes are often further modified to distinguish endogenous RNA from capped viral RNA as well (Ramanathan, Robb et al. 2016). Capping is also necessary for transcript splicing, as the cap binding complex of the splicing machinery must bind to it to initiate splicing (Konarska, Padgett et al. 1984, Ohno, Sakamoto et al. 1987). After processing, the cap is necessary for nuclear export and for translation initiation.

As the polymerase extends transcription, the splicing machinery will join with RNA pol II and initiate excision of introns. The cap-binding complex (CBC), consisting of a heterodimer of CBP80 and CBP20, binds to the 5' cap and recruits several small nuclear ribonucleoproteins (snRNPs) U4, U5, and U6 snRPs to initiate splicing. The spliceosome is a very large ribonucleoprotein complex that catalyzes the removal of the intron in a two-step process. Recognition of the 5' splice site is mediated by the snRNP U1 (Kondo, Oubridge et al. 2015). Cleavage of the RNA at the 5' splice site occurs first,

forming an intermediate lariat structure of the RNA. After cleavage of the 3' splice site and ligation of the remaining ends of the exon, the intron lariat is released. The spliceosome complex dissociates after each splicing event and re-associates downstream again for the next event.

Splicing not only removes introns from a transcript, but it also presents the opportunity for the cell to create more than one distinct mRNA from one gene via alternative splicing. Using alternative splice sites within the pre-mRNA, a message can be "adjusted" by exon skipping, intron inclusion, and extension or reduction of regulatory regions. This in turn can lead to changes in RNA stability, localization, and alter the resulting protein product. Alternative splicing is a ubiquitous event that allows for the creation of unique protein isoforms, and is essential for organismal development and responses to stimuli (Baralle and Giudice 2017).

The polyadenylation sequence (PAS) is a hexameric sequence occurring near the end of a gene that serves as a signal for transcription to end. The PAS is recognized by cleavage and polyadenylation specificity factor (CPSF), a multi-protein complex consisting of four main subunits. CPSF is recruited to the transcription initiation complex as transcription begins and follows the elongating polymerase (Dantonel, Murthy et al. 1997). The PAS sequence, AAUAAA, is specifically recognized by the CPSF30 subunit and cleavage of the transcript is initiated by CPSF73 (Ryan, Calvo et al. 2004, Mandel, Kaneko et al. 2006, Shimberg, Michalek et al. 2016). Poly-A polymerase (PAP) is recruited to the mRNA by an interaction with Fip1 (Factor interacting with Poly(A) polymerase), a subunit of CPSF that binds to U-rich elements upstream of the PAS (Kaufmann, Martin et al. 2004). Polyadenylation of the transcript results in approximately 200 to 300 adenosine nucleotides being added onto the end of the transcript (Wahle 1995, Kuhn, Gundel et al. 2009). The addition of a poly-A tail not only protects the transcript from exonuclease decay but serves a binding site for nuclear poly-A binding protein (PABPN1), that binds the tail in multimers and accompanies the mRNA out of the nucleus and into the cytoplasm.

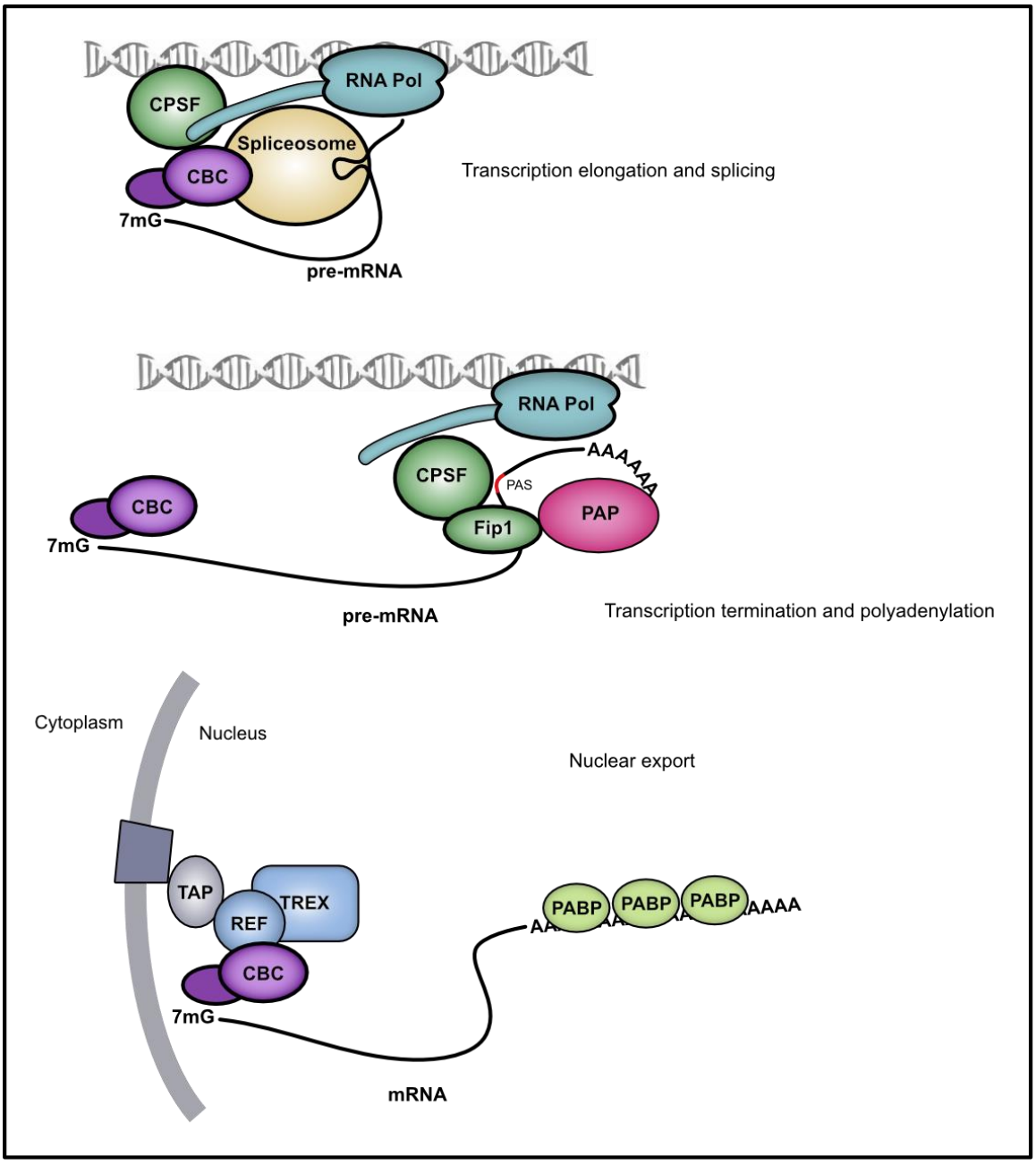


Figure 1.1: Pre-mRNA processing in the nucleus. RNA polymerase II (RNA pol) transcribes mRNA from DNA. A 5' cap (7mG) is added immediately to the transcript and bound by the cap-binding complex (CBC). As RNA pol II extends transcription, the spliceosome removes introns. Near the end of the gene, cleavage and polyadenylation specificity factor (CPSF) recognizes the polyadenylation sequence (PAS) cleaves the RNA and recruits poly(A) polymerase (PAP) to add non-templated adenosine nucleotides to the end of the transcript. Poly(A) binding protein (PABP) binds to the poly(A) tail. The CBC interacts with the REF subunit of the transcription-export complex (TREX), which will interact with the RNA export receptor (TAP) and facilitate transport of the mRNA out of the nucleus and into the cytoplasm.

Export out of the nucleus requires association of many cofactors, most of which accompany the transcript through transcription and pre-mRNA processing. The CBC binds to REF, an RNA-binding protein subunit of the transcription-export complex (TREX), which in turn will interact with the RNA export receptor, TAP. The spliceosome is also responsible for depositing cofactors along the RNA that are essential for its export through the nuclear pore, including REF (Le Hir, Izaurralde et al. 2000, Rodrigues, Rode et al. 2001, Iglesias and Stutz 2008, Gonatopoulos-Pournatzis and Cowling 2014).

1.2. Characteristics of mRNA

An RNA molecule is never “naked” in the cell. From the moment transcription begins until the time it is degraded, it is constantly in contact with other factors - the majority of these contacts are mediated through RNA binding proteins (RBPs). The processes of transcription, capping, splicing, and polyadenylation are all intricately tied together by protein complexes and the interactions between them. Each step in pre-mRNA processing can be said to serve as checkpoints, where certain proteins are deposited as marks that signal a successful processing event. This interplay ensures that complete and translation-competent mRNAs are exported to the cytoplasm.

The mature mRNA has several key features (Figure 1.2). In addition to the 5’ cap and poly(A) tail, the mature transcript also contains the coding sequence (CDS) that will be translated into the final protein and flanking untranslated regions (UTRs), which do not code for part of the protein. These UTRs serve as critical regulatory elements for post-transcriptional control.

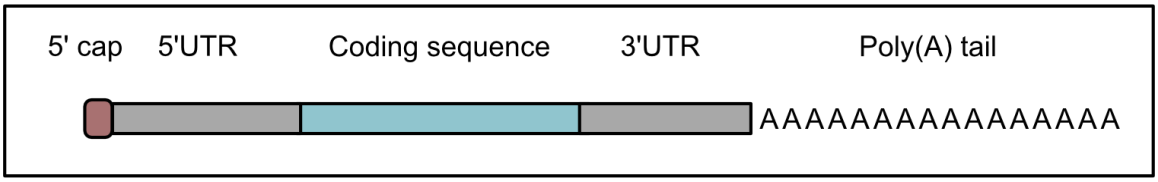


Figure 1.2: The general structure of a mature mRNA. The 5’ cap and poly(A) tail help protect the transcript from nucleolytic decay. The 5’ and 3’-UTRs are important regulatory regions. The coding sequence (CDS) is the only part of the RNA that will be translated into protein.

The 5’ cap protects the transcript from degradation in part by binding proteins such as CBC and PABP which compete with decapping enzymes for access to the cap structure. Removal of the cap is the first committed step toward transcript decay. In addition to its roles in pre-mRNA processing, the cap is also necessary for translation and translation-coupled quality control pathways.

The 5’-UTR of the transcript is the gateway to the coding region for the ribosome and many of the regulatory elements positioned here affect translation. In addition to sequences within the 5’-UTR that recruit RBPs, the RNA itself can form structures to inhibit or promote translation (Hinnebusch, Ivanov et al. 2016, Leppek, Das et al. 2018). One of the most important sequences in the 5’-UTR for translation initiation is the Kozak sequence, a consensus binding site adjacent to the CDS start codon. Alterations in the consensus can dramatically affect the amount of protein translated (Kozak 1986).

The coding sequence is delineated by start and stop codons, which determine the first amino acid and terminate translation, respectively. The start codon, AUG, codes for methionine and translation typically ends at the first stop codon, either UAA, UAG, or UGA. On some transcripts the ribosome does not stop at the first termination codon and continues translation in a process known as readthrough. Although readthrough is not common in eukaryotes, it can serve several purposes such as creating C-terminal extensions to proteins, suppressing premature stop codons, and incorporating additional signal sequences (Dabrowski, Bukowy-Bieryllo et al. 2015). In *Drosophila* an estimated 350 genes undergo readthrough and 42 readthrough events have been recorded in human cells (Lin, Carlson et al. 2007, Jungreis, Lin et al. 2011, Dunn, Foo et al. 2013).

The coding sequence itself can also serve as a regulatory mechanism. The degenerate protein code allows for several nucleotide triplets to code for the same amino acid. As such, there is the potential for some codons to be translated more efficiently than others based on the available tRNA pool, in a model known as codon optimality. This in turn can effect translation rates and RNA stability (Hanson and Coller 2018).

The 3’-UTR of a transcript is a molecular epicenter for regulatory motifs. Most miRNAs and RBPs bind in this region which is laden with both sequence and structural elements. The length of a 3’-UTR can be controlled by alternative polyadenylation (APA), a mechanism by which during transcription an alternative, usually upstream PAS,

is used to initiate cleavage and splicing. APA can be used in a tissue-specific manner, where the cell type expressing the transcript has need of different regulatory parameters. In addition to changing the length of the 3'-UTR, APA can result in the creation of proteins with different lengths or domains. (Edwalds-Gilbert, Veraldi et al. 1997, Lutz 2008). APA is also seen in cancer, where changes in 3'-UTR lengths effectively eliminate crucial regulatory elements that might otherwise illicit decay of the transcript (Mayr and Bartel 2009, Miles, Tschop et al. 2012). 3'-UTRs can also adopt secondary structures that may occlude binding sites for RBPs or miRNAs (Kedde, van Kouwenhove et al. 2010, Mayr 2017, Mayr 2019).

The length of the poly(A) tail added to a nascent transcript in the nucleus is controlled by the interaction between CPSF, PAP, and PABPN (Kuhn, Gundel et al. 2009). During the pioneer round of translation, nuclear PABPN is replaced by cytoplasmic PABPCs that cover the poly(A) tail with 25 nucleotide footprints (Smith, Gallie et al. 1997, Sato and Maquat 2009). PABPs have no enzymatic activity of their own but serve as binding sites for other proteins involved in translation, stability, and even decay. Since a long poly(A) tail coated with PABPC would theoretically inhibit the activity of a nuclease targeting the transcript, it has classically been believed that the longer a poly(A) tail was, the more stable the transcript would be. However, newer evidence suggests that transcripts with shorter tails are actively translated while longer tails are associated with lower abundance, poorly translated transcripts (Lima, Chipman et al. 2017, Nicholson and Pasquinelli 2019). Poly(A) tail length appears to govern translational status only during the very early stages of development (Eichhorn, Subtelny et al. 2016). Median tail length in the cytoplasm averages about 50 nucleotides in *Drosophila*, but poly(A) tail length is often dynamic and the mechanisms for controlling tail length are not well understood (Subtelny, Eichhorn et al. 2014).

The makeup of RBPs and other associated factors bound to an RNA ultimately determine its fate. The oversimplified central dogma that states that genetic information flows from DNA to RNA to protein neglects the myriad of other fates that await an RNA after leaving the nucleus, all of which can be mutually inclusive.

1.3. Cytoplasmic fates of mRNA

1.3.1. Translation

Translation of a transcript in the cytoplasm consists of multiple steps. The first is the formation and binding of the 43S preinitiation complex to the RNA. This is a complex consisting of multiple initiation factors including the CBC, eIF4F, and the 40S ribosome subunit. eIF4F binds the 5’ cap and unwinds the structure of the transcript to allow ribosome loading. PABP also binds to this complex, forming what is known as the closed loop model of translation, where the 3’ and 5’ ends of the RNA interact (Jackson, Hellen et al. 2010). It is likely that this interaction is driven in part by the inherent structure of the RNA itself, and natural dynamic 5’-3’ proximity (Vicens, Kieft et al. 2018). The 43S complex then scans the RNA in the 5’ to 3’ direction until reaching an initiation codon where translation will begin. Binding of the 60S ribosomal subunit and associated factors initiates translation.

Translation elongation and termination require the association of additional cofactors. Proteins known as elongation factors (eEFs) associate with amino-acyl tRNAs and are brought to the ribosome to join the matching amino acid to the growing peptide chain. Hydrolysis of GTP by eEF1A releases the tRNA into the aminoacyl site of the ribosome, while eEF2 promotes translocation of the tRNAs into the peptidyl and exit sites of the ribosome while the peptide bond is formed (Dever, Dinman et al. 2018). When the stop codon is reached, release factors eRF1 and eRF3 promote dissociation of the ribosome and the ribosomal subunits and cofactors are recycled for subsequent rounds of translation (Hellen 2018).

Translation is a highly regulated and dynamic process. It is at this step the subsequent amplification of gene product occurs. A single mRNA molecule can be translated many times, resulting in amplification of protein. On average, 2800 proteins are made from one mRNA (Schwanhausser, Busse et al. 2011). Translation can be controlled in multiple ways including regulating the activity of initiation factors via phosphorylation, controlling the protein levels of translation factors, altering the pool of tRNAs, and inhibiting necessary interactions between the initiation machinery and the ribosome. Many of the regulatory elements in a mRNA are positioned in the 3’-UTR. A

reasonable explanation for this is to allow RBPs to remain bound to the transcript and continue to exert their influence even while a transcript is actively being translated. An RBP bound to the 3'-UTR is typically inhibitory (with PABP being a notable exception). Many RBPs can stymie translation by interfering with the ability of eIF4F to bind to the cap. This is typically mediated through interactions with a family of proteins called 4E-BPs (eIF4E binding proteins) (Sonenberg and Hinnebusch 2009).

1.3.2. Storage and localization

An additional consideration of the total protein level derived from a transcript is its location. mRNAs often need to be transported to other cellular compartments for several reasons. Rather than rely on transport proteins or diffusion to relocate a finished protein to its final compartment, it is much more efficient to translate the protein near its ultimate site of action. Localization of transcripts can also prevent the accumulation of proteins in a given area, or enable rapid protein responses for functions such as signal transduction pathways (Blower 2013). Localization is critical during developmental processes, in neuron signaling, and co-translational insertion of endoplasmic reticulum (ER) proteins into the ER (Walter and Blobel 1981, Bashirullah, Cooperstock et al. 1998, Ohashi and Shiina 2020).

mRNAs can also be localized to sites of ribonucleoprotein (RNP) condensates such as processing bodies (p-bodies) and stress granules. It is thought that these membrane-less cytoplasmic granules are formed by liquid-liquid phase separation and driven by weak RNA-RNA and RNA-protein interactions, especially by aggregate-prone domains within the proteins (Mittag and Parker 2018). P bodies are typically enriched in proteins that function in the decay pathway, but it is not clear whether they are sites for RNA decay or sequestration and storage (Brengues, Teixeira et al. 2005, Luo, Na et al. 2018, Xing, Muhlrad et al. 2020). Stress granules contain translation initiation factors as well as decay proteins, and quickly form as a response to stress, initiating translational arrest (Kedersha, Chen et al. 2002, Markmiller, Soltanieh et al. 2018, Wolozin and Ivanov 2019). Upon relief of cellular stress the stress granules disassemble. Both stress granules and P bodies are highly dynamic and fluid objects and the mechanisms governing RNA entry and exit from them are still a matter of intense research.

1.3.3. Messenger RNA decay

The final fate of a mRNA is the decay pathway, which begins with removal of the poly(A) tail by a poly(A) ribonuclease. There are several ways in which removal of the poly(A) tail leads to the degradation of the transcript. In the absence of a poly(A) tail, PABP is unable to bind the mRNA and translation is inhibited. Deadenylation by the CCR4-Not (CNOT) complex is coupled to 5' cap removal and uncapped RNA in the cytoplasm are rapidly degraded by Xrn1, a non-specific 5' > 3' ribonuclease (Braun, Truffault et al. 2012). Transcripts can also be degraded in the 3' > 5' direction by the exosome.

In addition to the control of general mRNA levels, the decay machinery is also necessary for quality control. Translation of a transcript is monitored closely by cellular surveillance mechanisms to ensure product fidelity. Damaged transcripts, stable RNA secondary structures, or rare codons can lead to stalling of the ribosome. Translation stalling, or no-go decay, leads to endonucleolytic cleavage and subsequent degradation by Xrn1 and the exosome (Simms, Thomas et al. 2017). In a similar decay pathway, failure of the ribosome to terminate translation at the stop codon triggers non-stop decay (Kashima, Takahashi et al. 2014). Nonsense-mediated decay (NMD) occurs when the ribosome encounters a premature stop codon. In mammals, the activity of the spliceosome recruits the exon junction complex (EJC) to the RNA transcript. The EJC is a four-protein complex that binds to the RNA approximately 20 to 24 nucleotides upstream of the splice site and remains with the transcript until it is translated. The presence of the EJC downstream of a stop codon enhances NMD. The EJC does not seem to be necessary in *Drosophila*; rather the distance between the stop codon and the poly(A) tail determines whether the NMD pathway is initiated (Garneau, Wilusz et al. 2007, Choudhury, Singh et al. 2016).

There are multiple enzymes that can catalyze the removal of poly(A) tails in eukaryotes; the primary one in all organisms is the CNOT complex. Poly(A) ribonuclease (PARN) is also a deadenylase in mammals, but it is not present in *Drosophila*. Additionally, PARN has the ability to interact with the 5' cap, which can link translation directly to decay (Temme, Zaessinger et al. 2004, Godwin, Kojima et al. 2013). In *Drosophila*, almost all deadenylation is catalyzed by the CNOT complex and the PAN complex.

The PAN deadenylase complex is composed of the nuclease Pan2, and a homodimer of Pan3, which serves as an adaptor protein. PAN activity is dependent on PABP, an association mediated by an interaction between PABP and Pan3. The PAN complex is poly(A) specific and distributive, and processes poly(A) tails until approximately 25 nucleotides remain, believed to be the footprint of a single PABP (Wahle and Winkler 2013). The PAN complex appears to function in concert with the CNOT complex on certain transcripts, with the PAN complex initiating decay and the CNOT complex completing poly(A) tail removal (Yi, Park et al. 2018).

The CNOT complex is a megadalton cytoplasmic complex conserved in all eukaryotes (Figure 1.3). In *Drosophila*, it consists of 5 to 6 stably associated subunits: Not1, Not2, Not3, Caf40 (Not9 in mammals), CCR4, and Pop2. Not4, an E3 ubiquitin ligase, is not always detected in immunoprecipitation assays and is not considered to be a stable complex member. Two accessory proteins, Not10 and Not11, can be detected with the CNOT complex as well (Temme, Zhang et al. 2010, Wahle and Winkler 2013).

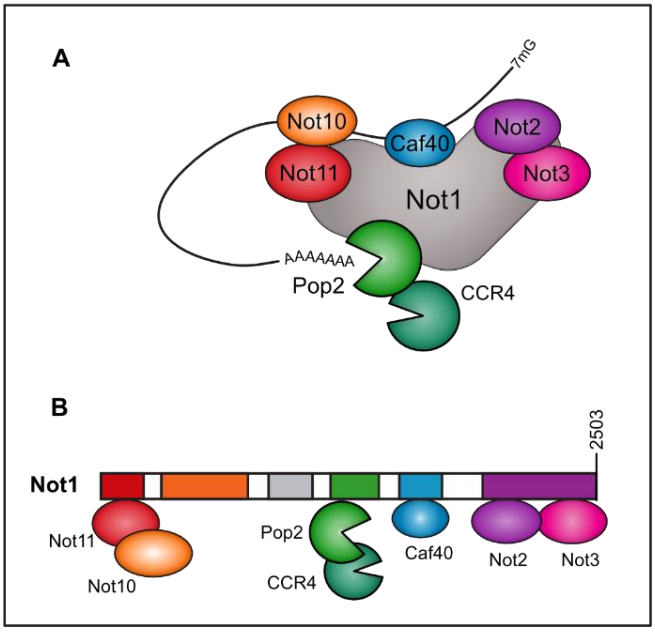


Figure 1.3: The CNOT deadenylase complex. (A) Depiction of assembled complex. Not1 is the largest member and acts as a scaffold for the other subunits. Caf40 has non-specific RNA-binding activity and Pop2 is the active deadenylase in Drosophila. (B) Linear representation of Not1 and the modules where the subunits bind. Figure adapted from (Chang, Muthukumar et al. 2019).

Pop2 (also called CAF1) and CCR4 are the two deadenylase enzymes associated with the complex. CCR4 is an exonuclease/ endonuclease/ phosphatase (EEP)- family deadenylase and is the primary catalytic deadenylase in budding yeast, but it appears to have no activity in *Drosophila*. Pop2, a DEDD-class nuclease, is the primary catalytic subunit in *Drosophila* (Temme, Zaessinger et al. 2004). Each deadenylase has two orthologs in mammals: CNOT6 and CNOT6L for CCR4, and CNOT7 and CNOT8 for Pop2. The association between CNOT6/6L with the CNOT complex is mutually exclusive, as it is for CNOT7/8. Both CNOT6/6L and CNOT7/8 are catalytically active in mammalian cells.

Not1 is the largest protein of the complex and forms the scaffold that binds the other subunits. In *Drosophila*, Not1 is approximately 2500 amino acids long and 280 kDa. Not2 and Not3 both bind Not1's C-terminal domain and together comprise the core NOT module. In *Drosophila*, only the NOT module and Pop2 are required for deadenylase activity (Temme, Zhang et al. 2010), however *in vitro* other human Not complex members can stimulate activity (Raisch, Chang et al. 2019).

The CNOT complex links deadenylation to decapping and other decay pathways through multiple interactions with cofactors. The decapping complex, composed of Dcp1, Dcp2, and accompanied by decapping enhancers (Edcs), binds to Not1 through an interaction with Edc3 and Me31b (DDX6 in humans) (Chen, Boland et al. 2014). Once a transcript is decapped, it is rapidly degraded by Xrn1. This is ensured via a direct interaction between Xrn1 and Dcp1 (Braun, Truffault et al. 2012) and a direct interaction between Xrn1 and the CNOT complex (Chang, Muthukumar et al. 2019).

Both the PAN and CNOT complexes can be recruited to the mRNA in a non-specific manner by their interactions with PABP. For CNOT, the interaction is bridged by another protein, Tob (Funakoshi, Doi et al. 2007). The CNOT complex member Caf40 also has intrinsic RNA binding activity (Garces, Gillon et al. 2007, Raisch, Chang et al. 2019). Targeted recruitment of the deadenylase complex to a transcript can occur through interactions with sequence specific RNA binding proteins, or by microRNAs (miRNAs).

miRNAs are ~22 nucleotide-long single-stranded RNAs loaded onto Argonaute proteins that can base pair with sequences primarily in a transcript's 3'-UTR. Argonaute proteins are endonucleases, however the miRNA they carry must have near-perfect

complementarity to the mRNA target in order for them to cleave the RNA (Elbashir, Martinez et al. 2001, Hutvagner and Zamore 2002, Bartel 2009). Since this is rarely the case in metazoans, Argonautes and miRNAs mediate gene silencing by recruiting the decay machinery via interactions with GW182 proteins (TNRC6A/B/C in vertebrates). GW182 proteins contain N-terminal glycine and tryptophan repeats that interact with Argonaute proteins, and C-terminal tryptophan repeats that can bind either the CNOT or PAN complexes (Behm-Ansmant, Rehwinkel et al. 2006, Braun, Huntzinger et al. 2011, Fabian and Sonenberg 2012, Chen, Boland et al. 2014, Gebert and MacRae 2019).

The CNOT complex interacts with multiple sequence-specific RNA-binding proteins to cause deadenylation of targeted transcripts. The TNF-α mRNA regulator, Roquin, binds to a conserved stem loop structure in the 3'-UTR and recruits the CNOT complex via an interaction with CNOT 9 (Caf40) (Leppek, Schott et al. 2013, Tan, Zhou et al. 2014, Sgromo, Raisch et al. 2017). *Drosophila* Bag of marbles (BAM) also recruits the CNOT complex through Caf40 (Sgromo, Raisch et al. 2018). Tristetraprolin (TTP; Tis11 in *Drosophila*) is an RBP that binds to A-U rich sequences. TTP interacts directly with Not1 through multiple contacts (Sandler, Kreth et al. 2011, Fabian, Frank et al. 2013). Nanos, an important developmental morphogen, contains an N-terminal Not1 interaction motif (NIM) with which it binds to the Not module (Bhandari, Raisch et al. 2014). Interestingly, *Drosophila* Nanos does not contain the same motif but it retains the ability to contact the Not module including Not3 through an extended Nanos effector domain (NED) (Raisch, Bhandari et al. 2016). Most importantly, the PUF (Pumilio and FBF) family of RNA binding proteins direct regulation of their targets by recruiting the decay machinery.

1.4. Pumilio and PUF proteins

PUF proteins are defined by a characteristic RNA binding domain (RBD), termed the Pum-homology domain (Zamore, Williamson et al. 1997). PUF proteins bind specifically to sequences called Pumilio response elements (PREs). The founding members of the PUF family of proteins, *Drosophila* Pumilio (Pum), and *C. elegans fem-3* binding factor (FBF) were discovered by their ability to bind and regulate target mRNAs

during development. It is these two proteins that give the family their name. The PUF family is conserved in plants and throughout eukaryotes, though the more closely related proteins in higher organisms are referred to as Pum proteins. *Drosophila* has a single Pum protein, whereas mammals have two, PUM1 and PUM2.

Since their initial discovery, many targets of Pum proteins have become known, and their roles in important biological functions are beginning to be elucidated. To date, Pums have been implicated not only in development, but in cancer (Miles, Tschop et al. 2012, Naudin, Hattabi et al. 2017, Brocard, Khasnis et al. 2018), germ line maintenance (Lin and Spradling 1997, Forbes and Lehmann 1998, Crittenden, Bernstein et al. 2002, Chen, Zheng et al. 2012), learning and memory formation (Dubnau, Chiang et al. 2003, Baines 2005, Chen, Li et al. 2008), neurogenesis (Fiore, Khudayberdiev et al. 2009, Zhang, Chen et al. 2017, Martinez, Randolph et al. 2019), aging (Kopp, Elguindy et al. 2019), cell cycle control (Lin, Qiang et al. 2019), mitochondrial homeostasis (D'Amico, Mottis et al. 2019), neurodegenerative disease (Gennarino, Singh et al. 2015, Gennarino, Palmer et al. 2018), cellular signaling (Bermudez, Jouandin et al. 2011), and immune response (Narita, Takahasi et al. 2014). RIP-chip (RNA immunoprecipitation-microarray analysis) data in both human and *Drosophila* indicate that Pum proteins could potentially regulate 15% of the cell’s total transcriptome (Galgano, Forrer et al. 2008).

1.4.1. Essential roles of Pum proteins

Pum proteins are essential in higher organisms. Pum mutant flies fail to develop past the larval stage (Nusslein-Volhard, Frohnhofer et al. 1987). In mammals, double knockout Pum mice result in embryonic failure shortly after the blastocyst stage (Lin, Zhang et al. 2018).

The first identified Pum targets

In *Drosophila*, Pum (along with Nanos and Brat) is responsible for the correct abdominal segmentation in the embryo by regulating hunchback mRNA, a transcription factor that controls patterning genes. Loss of Pum results in improper accumulation of hunchback protein in the posterior and embryos fail to develop abdominal segments

(Barker, Wang et al. 1992, Murata and Wharton 1995). In *C. elegans*, the *fem-3* mRNA is regulated by FBF to mediate the sperm to oocyte switch. *C. elegans* hermaphrodites produce sperm in a larval stage and transition to oocyte production as adults. During sperm production, TRA-1, a transcriptional regulator of feminizing genes, is suppressed by FEM-3. For the sperm to oocyte switch to occur, FEM-3 must be “turned off”. This is accomplished by decay of *fem-3* mRNA by FBF proteins (Ahringer, Rosenquist et al. 1992, Zarkower and Hodgkin 1992, Zhang, Gallegos et al. 1997, Bachorik and Kimble 2005).

Developmental and signaling roles

Wing vein development in *Drosophila* is a process under the control of several different signaling pathways including epidermal growth factor receptor (EGFR) and Notch (Johannes and Preiss 2002). During the larval wing cell differentiation stage, enhanced EGFR signaling results in the generation of extra wing veins, a morphology phenocopied in Pum mutants (Kim, Kim et al. 2012). Yeast three-hybrid screens identified several RNAs in the EGFR pathway as Pum targets: egfr, rolled (human homolog MAPK1), the guanine exchange factor sos, and the adapter protein drk (human homolog GRB2). In a follow up luciferase reporter assay, overexpression of Pum caused repression of these targets, but not when the Pum binding sites were mutated (Johannes and Preiss 2002).

In *Drosophila*, Pum proteins are necessary to keep migrating germline progenitor cells in a state of mitotic arrest. Pum mutants commence mitosis early and fail to migrate properly which ultimately leads to germline failure (Asaoka-Taguchi, Yamada et al. 1999). Pum exhibits this control over the cell cycle by regulating cyclinB RNA in partnership with Nanos. Interestingly, Pum’s role in this regard appears to be for targeting, since Nanos is capable of repressing cyclinB RNA without Pum if it is tethered, and Pum does not bind to this transcript in the absence of Nanos (Kadyrova, Habara et al. 2007, Weidmann, Qiu et al. 2016).

In human B cells infected with Epstein-Barr virus (EBV), Pums contribute to the control of poly(A) tail length of another important cell cycle protein, RGCC (Brocard, Khasnis et al. 2018). Expression of RGCC is essential for the growth of cells infected

with EBV. Depletion of Pum proteins allow an increase in protein levels of RGCC and contributes to the proliferation of infected cells.

Several RIP-chip analyses have identified another cell cycle regulator, p27 (gene CDKN1B), as a bound target of both mouse and human PUM1 and PUM2 (Galgano, Forrer et al. 2008, Zhang, Chen et al. 2017). In human fibroblasts, knockdown of PUM1 causes an increase in the levels of p27 protein (Kedde, van Kouwenhove et al. 2010). In addition to increased p27 protein levels, depletion of human PUM1 or PUM2 halted cell cycle progression and caused a delay of entry into S phase when cells were treated with growth factor.

Neuronal roles

Pum proteins have multiple roles in the *Drosophila* brain and neuromuscular system. In the peripheral nervous system, over-expression of Pum (and Nanos) reduces the number of high-order dendritic branches in class III and IV neurons (Ye, Petritsch et al. 2004). Overexpression of Pum during development of the neuromuscular system causes errors in axon guidance and leads to truncations in synapses (Kraut, Menon et al. 2001). At the neuromuscular junction (NMJ) Pum is necessary in neurons for proper bouton size and number, and in muscle Pum regulates the RNA of the glutamate receptor GluRIIa and controls translation of many other genes by targeting translation initiation factor eIF4E RNA. Reduction of Pum in this system causes accumulation of eIF4E aggregates. Additionally, Pum mutants have a higher frequency of spontaneous neurotransmitter release (Menon, Sanyal et al. 2004, Menon, Andrews et al. 2009).

Neuronal hyperexcitability, observed as prolonged releases in neurotransmitters, is seen in Pum mutants, whereas over-expression of Pum can prevent such hyperexcitability in the NMJ (Schweers, Walters et al. 2002). This is due to the regulation of *paralytic*, which encodes the sole *Drosophila* voltage gated sodium ion channel. Over-expression of Pum protein causes a decrease in the levels of paralytic mRNA. Conversely, Pum mutants show two-fold increases in the levels of paralytic (Mee, Pym et al. 2004).

The formation of long-term memory relies heavily on localized translation. RNA packaged into neural RNP granules for transport remains in a quiescent state until it is

delivered to the active synapse, where translational de-repression commences. Pum was implicated as one of the proteins responsible for repressing translation in neural granules (Dubnau, Chiang et al. 2003). In a screen for memory formation mutants in *Drosophila*, two different p-element mutations of Pum were identified. These mutants show severe defects in one day memory after spaced training. Additionally, levels of Pum transcripts were found to increase during memory formation (Dubnau, Chiang et al. 2003).

1.4.2. Regulation of Pum

1.4.2.1. Regulation of Pum RNA

Little is known about how Pum itself is regulated, although Pum RNA is likely auto-regulated by the presence of 2 perfect PREs in the 3'-UTR. In humans, both PUM1 and PUM2 RNA possess PREs. Pum RNA is frequently identified in RIP-Chip experiments as targets of Pum proteins, in *Drosophila*, mouse, and humans (Gerber, Luschnig et al. 2006, Galgano, Forrer et al. 2008, Morris, Mukherjee et al. 2008, Chen, Zheng et al. 2012, Laver, Li et al. 2015, Zhang, Chen et al. 2017). Via alternative polyadenylation, *Drosophila* Pum transcripts can be made with three different lengths of 3'-UTRs, allowing for varying levels of regulation determined by developmental stage or other factors.

Embryonic lethal abnormal visual system (ELAV), an RBP involved in alternative polyadenylation in developing neurons, binds to Pum RNA (Hilgers, Lemke et al. 2012). By associating with the longer form of Pum RNA near a proximal polyadenylation signal, ELAV likely regulates the production of longer Pum mRNA isoforms. In this model, ELAV binding to the proximal PAS site promotes transcriptional readthrough and results in the longer form 3'-UTR. Since ELAV itself is specific to neural tissues and is both temporally and spatially regulated, this model could describe one of many ways in which Pum activity is managed in developing neurons.

Pum RNA is regulated by the RBP RbFox1 protein during *Drosophila* germline development. The 3'-UTR contains four Rbfox1 binding sites, two of which are conserved across many *Drosophila* species. Loss of Rbfox1 resulted in increased levels of pum RNA and germline tumor formation (Carreira-Rosario, Bhargava et al. 2016). In

addition, mouse Pum1 and Pum2 were identified as Rbfox targets by cross-linking and immunoprecipitation followed by high-throughput sequencing (CLIP-seq) in mouse neurons (Weyn-Vanhentenryck, Mele et al. 2014).

A 2016 CLIP-seq experiment in S2 cells identified Pum RNA as a target of Orb2, the *Drosophila* Cytoplasmic polyadenylation element binding protein 2 (CPEB2) homolog. The shortest form of the Pum 3'-UTR contains 2 Orb2 binding sites, UUUGU, while the longer transcript variants contain a total of 9 Orb2 binding sites. Using a luciferase reporter containing the Pum 3'-UTRs, the study showed that Pum RNA is repressed by overexpression of Orb2 when compared to overexpression of an Orb2 RNA-binding mutant (Stepien, Oppitz et al. 2016).

In *Drosophila* ovaries, Pum mRNA was also identified in a RIP-seq experiment for Hrp38, the fly homolog of hnRNP A1. This was further validated by RIP-RT-PCR. Hrp38, best known for its roles in splicing and translation during development, also targets and represses Nanos RNA (Borah, Wong et al. 2009, Ji and Tulin 2016).

Pum RNAs are also targets for miRNA-induced decay. Mammalian Pum2 is a direct target of miR134 in hippocampal neurons, where depletion of Pum RNA is necessary to promote dendritic outgrowth (Fiore, Khudayberdiev et al. 2009, Fiore, Rajman et al. 2014). Although there are no studies directly linking *Drosophila* miRNA to Pum regulation, TargetScan lists several potential miRNA seed sites in the 3'-UTR of the transcript, one of which matches the conserved family precursor miR 10.

Given all of the crucial roles Pum proteins have in development and overall homeostasis, it is likely that Pum RNA is regulated by many other RBPs as well.

1.4.2.2. Regulation of Pum protein

Pum could potentially be regulated by post-translational modifications, including phosphorylation. These modifications could affect Pum's target specificity, auto-regulation, or binding partners. During the egg activation stage of embryonic development, Pum is regulated by the Pan Gu (PNG) kinase complex, a master kinase of oocyte to embryo transition. This model describes a sequence of events in which Pum is bound to cyclin B RNA and thus represses its translation, until phosphorylation by PNG triggers de-repression of the transcript (Vardy and Orr-Weaver 2007). Although cyclin B is a well-characterized co-target of Pum and Nanos, the effects of

phosphorylation on Pum have not been studied. I*n vitro*, a GST-Pum fusion protein is a substrate for PNG, but it is unclear whether this interaction occurs *in vivo*, or if it is biologically relevant (Hara, Lourido et al. 2018). In human fibroblasts, phosphorylation of PUM1 at serine 714 is rapidly increased following epidermal growth factor stimulation, and a phosphomimetic mutation (S714E) increased PUM1's binding activity (Kedde, van Kouwenhove et al. 2010). Similarly, phosphopeptide enrichment in *Drosophila* embryos followed by mass spectrometry identified five phosphorylated serine residues (Zhai, Villen et al. 2008). Phoshopep lists an additional two observed phosphorylated Pum peptides, but the kinase(s) responsible for any of them have not been identified (Bodenmiller, Malmstrom et al. 2007). Another study in human HEK293 cells indicated that PUM2 is phosphorylated by the serine/threonine kinase Aurora A, and then physically associates with Aurora A to stabilize it against proteasomal degradation. This in turn leads to the accumulation of Aurora A protein and kinase activity, which allows the cell to enter mitosis (Huang, Wu et al. 2011). Although this model makes no mention of the function of Pums as RNA decay factors, the implication of Pums having other roles is intriguing.

Long noncoding RNAs (lncRNAs) are reported to function as microRNA sponges to affect the available pool of miRNAs and thus counteract their repressive activities. Non-coding RNA activated by DNA damage (NORAD), a lncRNA with 17 PREs, functions in this manner against Pum proteins to regulate expression of Pum targets (Tichon, Gil et al. 2016). NORAD expression increases upon DNA damage and loss of NORAD leads to genomic instability (Lee, Kopp et al. 2016). Presumably this effect is due to increased activity of Pum in targeting specific DNA damage response factors such as centromere components, DNA replication factors, and proteins involved in mitosis (Lee, Kopp et al. 2016). NORAD is highly conserved, but only in mammals. In addition to NORAD, HCG11, a pseudogene of NORAD, also contains 17 PREs, although its effects on Pum are not known and its expression levels are reported to be much lower than NORAD (Chen, Bao et al. 2019). However, in lung cancer patients, both NORAD and HCG11 underexpression are associated with poor patient prognosis (Stewart, Enfield et al. 2019). It is also unclear if NORAD or HCG11 exist at high enough levels to affect the cellular amount of Pum proteins, but under select conditions it could be possible.

1.4.3. Protein interactors of Pum

Pum's mechanism of repressing target RNAs is multi-faceted and mediated through interactions with protein partners. During embryonic development, Pum associates with the RBP Nanos to control hunchback RNA (Arvola, Weidmann et al. 2017). A consequence of Nanos binding cooperatively with Pum is enhancement of Pum binding and a potential increase of the target pool (Weidmann, Qiu et al. 2016). Additionally, the recruitment of Nanos is essential for binding certain targets like cyclin B, for which Pum alone has no affinity. A partnership with Nanos might also recruit the CNOT complex through Nanos's NIM or NED. In addition to Pum and Nanos, the RBP Brat also binds and regulates hunchback mRNA, though it remains unclear if Brat interacts with Pum or Nanos, or modulates their repressive activity (Arvola, Weidmann et al. 2017).

In *Xenopus*, Pum interacts with cytoplasmic polyadenylation element binding protein (CPEB), a protein responsible for initiating cytoplasmic polyadenylation during oocyte maturation. CPEB interacts with poly(A) polymerase (PAP) to lengthen transcript poly(A) tails and allow for translation (Richter and Lasko 2011). Depending on the context of CPEB binding sites in the 3'-UTR, Pum's association with CPEB can either enhance translation or enact repression (Pique, Lopez et al. 2008). Evidence exists of a conserved interface between PUF proteins and CPEBs, and human PUM2 was demonstrated to interact with CPEB3 via this interface (Benoit, Papin et al. 2008, Campbell, Menichelli et al. 2012).

C. elegans FBF and human Pum2 interact with AGO proteins and disrupt translation elongation by inhibiting the GTPase, eEF1a (Friend, Campbell et al. 2012). Although the RBDs of Pum2 and *Drosophila* Pum interact with AGO *in vivo*, this association is not necessary for translational repression. Rather, repression occurs through antagonism of PABP, although the exact mechanism is not clear. Deadenylation is enhanced by the RBD, presumably by the recruitment of the CNOT complex (Weidmann, Raynard et al. 2014).

Mass-spectrometry of human Pum complexes revealed the presence of multiple members of the CNOT complex, and subsequent analysis using immunoprecipitation assays showed physical associations with the deadenylase subunits *in vivo* (Van Etten, Schagat et al. 2012). *In vitro*, both the N-terminus and the RBD of *Drosophila* and

human Pums are able to interact with the CNOT complex (Arvola, Chang et al. 2020, Enwerem, Elrod et al. 2021). Additionally, the ability of Pums to repress is dependent on the deadenylase activity of the CNOT complex although the mechanism of this interaction is not clear.

1.4.4. The structure of Pum proteins

Between the human and fly proteins, there is 80% identity within the RNA-binding domain (RBD) (Goldstrohm, Hall et al. 2018). The RBD, like all PUF family members, is composed of 8 *alpha*-helical repeats. Each repeat consists of ~40 amino acids arranged in 3 helix-turn motifs. The overall structure adopts a crescent shape, where the concave face binds the RNA using aromatic and basic residues (Wang, Zamore et al. 2001). Base-staking and Watson-Crick interactions with the RNA give Pum precise sequence specificity to the nucleotides UGUANAUA, with nanomolar affinity (Weidmann, Qiu et al. 2016). The Pum RBD binds to the RNA in a “backward” orientation, where the C-terminus of the protein faces the 5’ end of the transcript. The eighth *alpha*-helical repeat binds the first nucleotide of the RNA sequence.

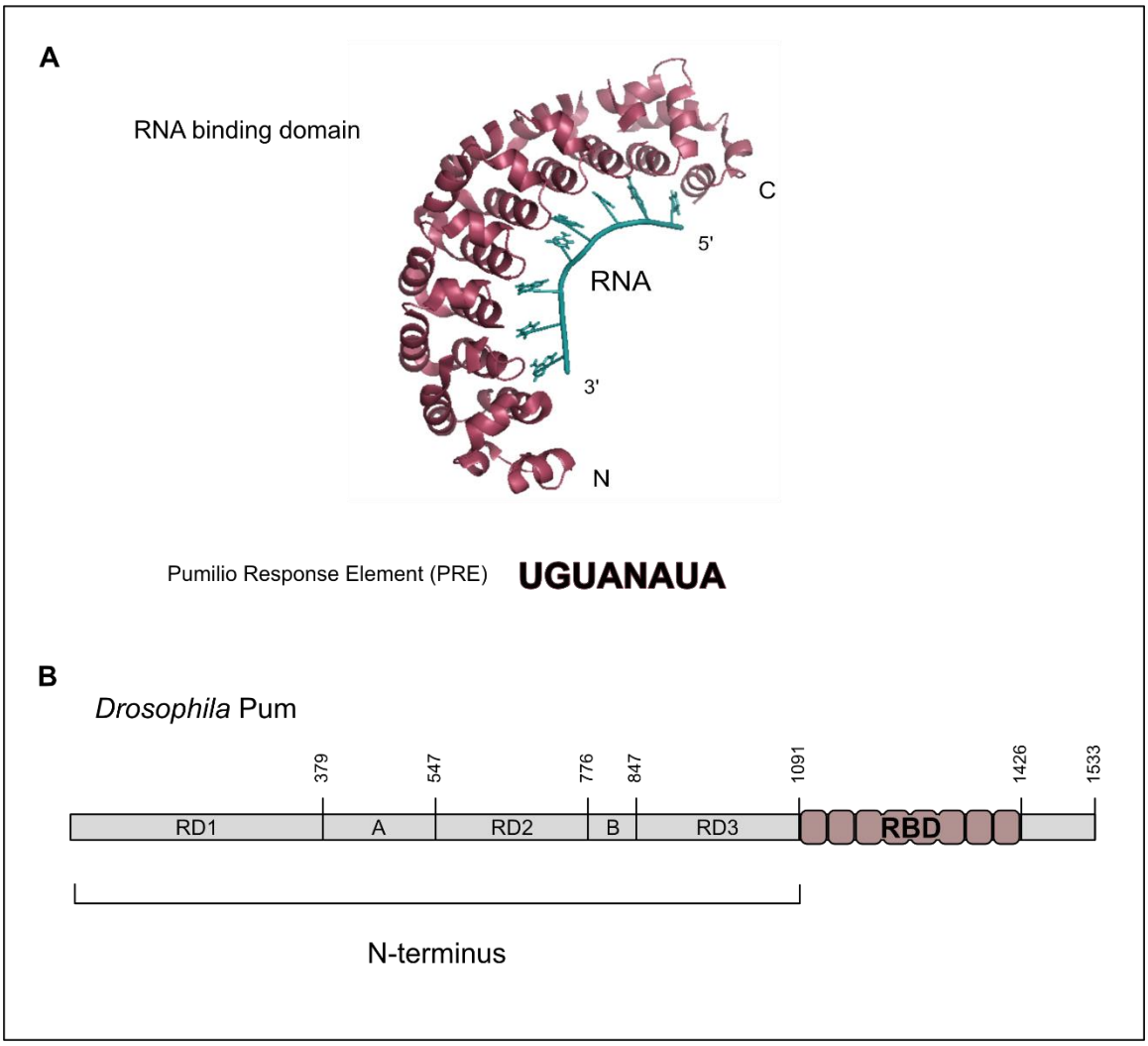


Figure 1.4: The structure of Pum. (A) Crystal structure of the RNA binding domain, PDB entry 5KLA; image generated using PyMOL. Below the structure is shown the PRE sequence that Pum binds to. (B) Schematic representation of *Drosophila* Pum protein, showing repression domains (RDs), PCMa and PCMb, and the RBD. Amino acid positions are shown above.

Although the RBD is responsible for targeting Pum to a transcript, the RBD alone is not sufficient to rescue the segmentation defect of Pum-null *Drosophila* embryos (Wharton, Sonoda et al. 1998). In the *Drosophila* NMJ, full-length Pum is necessary to rescue the neuromuscular defects seen in Pum mutants, and regulation of voltage gated sodium channels and *paralytic* RNA requires full-length Pum (Menon, Sanyal et al. 2004, Muraro, Weston et al. 2008). In addition, in a luciferase assay reporter system, the RBD alone accounted for only 22% of the repressive activity when compared to the full-length

protein, indicating that the majority of Pum's repressive ability is contained in the N-terminus (Weidmann and Goldstrohm 2012).

The N-terminus of Pum is predicted to be disordered, and no crystal structures have been solved. The N-terminus is divergent between human and *Drosophila*, however two regions of conserved residues exist, PCMa and PCMb (Weidmann and Goldstrohm 2012). PCMa, located at residues 379 to 547, shares 38% identity with human Pums, and PCMb, residues 777 to 847, shares an average of 65%. The intervening areas between PCMa and PCMb of the N-terminus can be divided into three unique repressive domains (RDs): RD1 (amino acids 1 to 379), RD2 (aa 547 to 777), and RD3 (aa 847 to 1090). In a tethered luciferase assay reporter system, each individual RD maintains the ability to affect decay, with RD2 and RD3 having the highest levels of activity (Weidmann and Goldstrohm 2012). The PCMs have little to no repressive activity on their own and appear to have auto-regulatory functions. Expressing PCMb as a fusion construct to the RDs can counteract the repression conferred by an RD, and PCMa appears to counteract PCMb. It is not clear how these disparate regions of Pum function together to bring about target mRNA decay.

Much has been learned about Pum proteins since their discovery over thirty years ago, but many questions remain. One aspect yet to be fully explored is the mechanism by which Pums cause decay of target transcripts. Pum proteins effect decay through an association with the CNOT complex, but the residues or regions responsible have not been identified. In this work we seek to understand how RD3 contributes to the function of Pum in *Drosophila*.

Chapter 2: Structure-function analysis of RD3

2.1. Introduction

Pum enacts decay by recruiting the CNOT deadenylase complex. Both *Drosophila* and human Pums interact with the complex (Van Etten, Schagat et al. 2012, Arvola, Chang et al. 2020, Enwerem, Elrod et al. 2021) and Pum's activity is severely diminished when Not1 or Pop2 are depleted (Arvola, Chang et al. 2020).

We hypothesized that Pum directly recruits the CNOT complex through short interaction motifs. We sought to identify those regions within Pum that are responsible for recruiting the CNOT complex and to define which members of the complex were necessary for the interaction.

The functional redundancy of human PUM1 and PUM2 makes it difficult to evaluate their mechanism of action. With only a single Pum protein, *Drosophila* Pum is well-suited to functional and biochemical characterization. *Drosophila* cells are simple to maintain in culture, and highly amenable to manipulation. We employed this system to interrogate the method of Pum repression.

The three repression domains (RDs) in the N-terminus of Pum exhibit functional redundancy and removal of a single RD has little effect on the activity of the protein (Weidmann and Goldstrohm 2012). Because of this, it was necessary to focus on a single RD in order to examine the mechanism of repressive activity. RD3 was chosen for several reasons. First, in human Pums, RD3 contains the most activity (Weidmann and Goldstrohm 2012). Human RD3 also interacts with the CNOT complex *in vitro*, whereas RD1 and RD2 do not (Enwerem, Elrod et al. 2021). Second, RD3 is the only RD present in all annotated protein isoforms. The shortest isoforms in *Drosophila* contain only a small portion of RD2, followed by PCMb, RD3, and the RBD. The spatial and temporal expression levels of the different Pum isoforms have not been studied in detail. It is therefore possible that in certain developmental contexts, RD3 will have the highest biological significance. Third, RD3 is the most conserved of the three RDs. Human PUM2 does not have a functional RD1, and RD2 has 16% amino acid identity between *Drosophila* and human Pums, whereas RD3 has 23% identity.

In *Drosophila*, the Pum gene has nine annotated transcripts (A-H, illustrated in Figure 2.1) and encodes five unique protein isoforms. Isoform A is considered the

canonical protein and is 1533 amino acids in length. This form is approximately 160 kDa in size (Thurmond J 2019). The shorter isoforms are predicted to make proteins around 98 kDa (isoforms E, F, and H) or 130 kDa in size (isoform B). In adult flies, ovaries and embryos, both the 160 and 130 kDa isoforms are present (Parisi and Lin 1999). It is not known under which conditions the shorter isoforms are expressed. Dmel2 cells express both 130 and 160 kDa forms, but only the 160 kDa form is detected in Dl1 cell lines via western blot analysis.

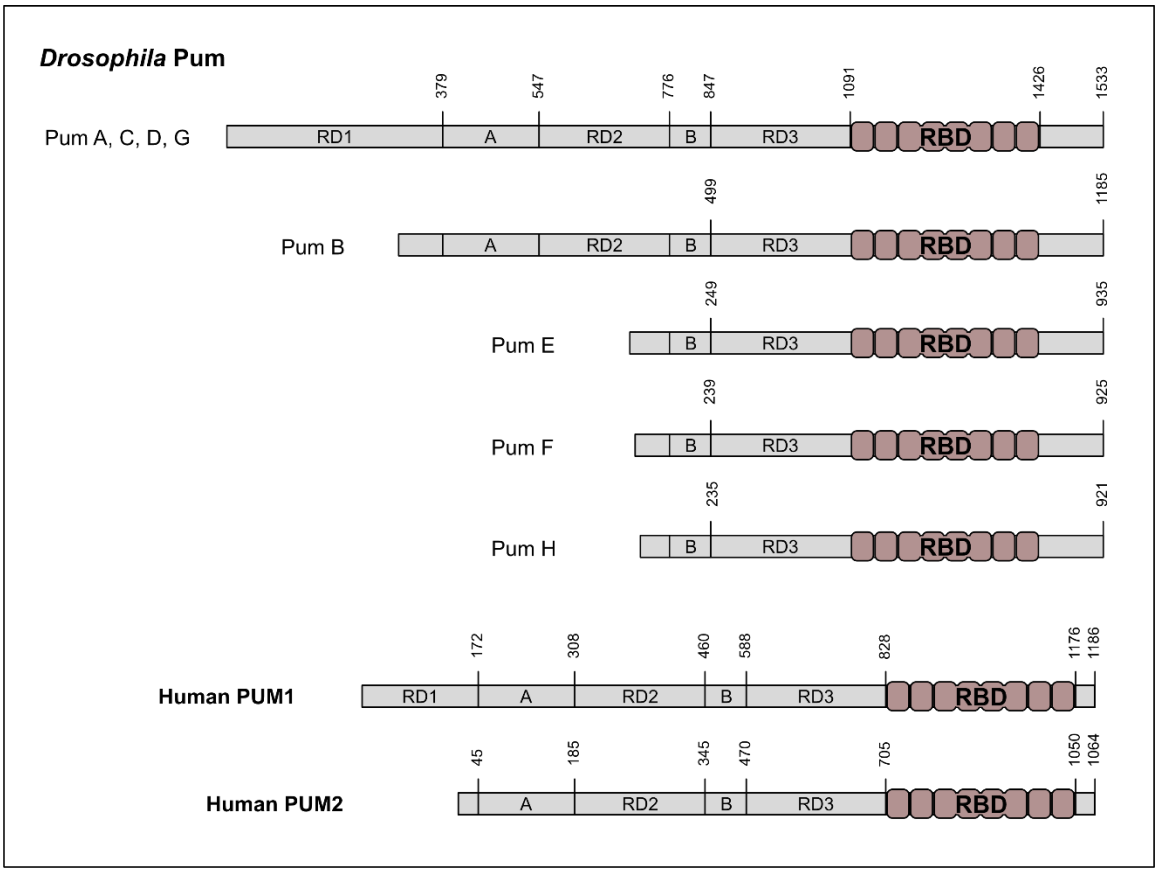


Figure 2.1: *Drosophila* Pum isoforms. Isoform A is considered the canonical isoform. The transcripts that encode isoforms A, C, D, and G differ in their UTRs. Note that in isoforms E, F, and H, RD3 is the only complete RD. Human PUM1 and PUM2 are also shown for comparison. *Drosophila* Pum is most closely related to human PUM2, despite PUM2's lack of RD1.

2.2. Overall structure of RD3

In *Drosophila*, RD3 is comprised of amino acids 847 to 1090 (Figure 2.1). This area lies between PCMb on the 5' end and is immediately followed by the RBD (Weidmann and Goldstrohm 2012). RD3 contains many stretches of sequence dominated by a single amino acid, specifically alanine and/or glutamine. It has been proposed that similar stretches in other proteins might play a role in forming protein aggregates or serve as spacer regions between functional motifs (Laughon, Carroll et al. 1985, Macdonald 1992). For RD3, these intervening regions show very little conservation, even in other insect species. The entire Pum N-terminus contains numerous low complexity areas that are not amenable to structural analysis. For this reason, our approaches were limited to functional analyses of the conserved regions.

2.3. Identification of conserved residues in RD3

Using an alignment of 82 Pum proteins across 45 different species, we sought to identify regions of conservation that could indicate functional residues. The alignment was generated with Clustal Omega and ConSurf using Pum proteins from species including insects, reptiles, birds, and mammals. This alignment highlighted several interesting features of Pum proteins, including several conserved glutamine-rich regions and multiple regions of alanine stretches present in insects. From this alignment, five regions of conservation were annotated in RD3: Q-P/N-rich region, CR1, CR2, CR3, and CR4 (Figure 2.2).

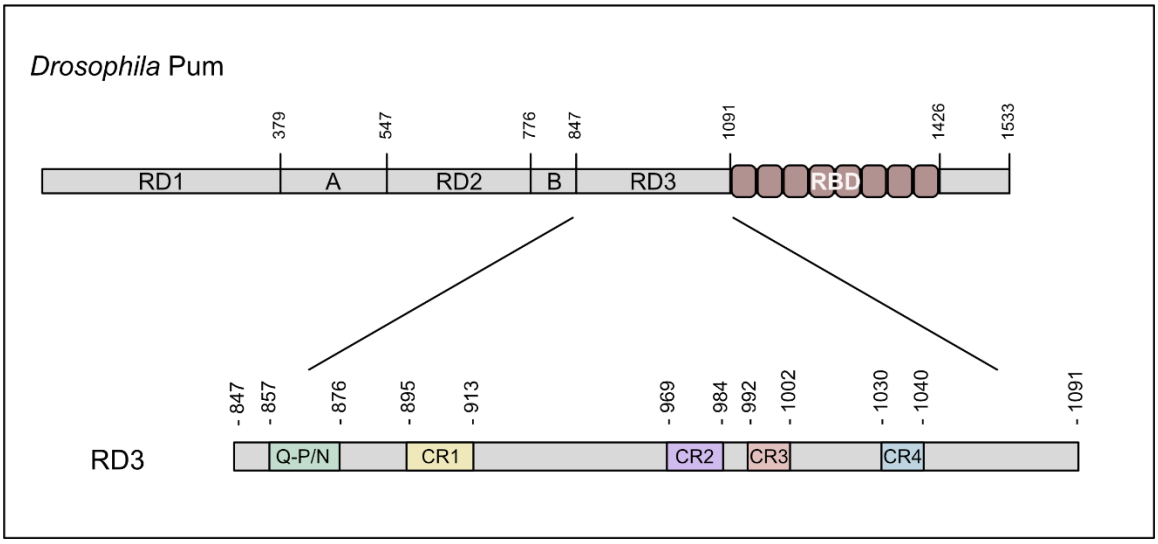


Figure 2.2: Location of conserved regions (CRs) in RD3 of *Drosophila* Pum. Full-length Pum is shown at top, with an expanded view for RD3 below. The amino acid positions for each CR within RD3 are indicated.

The Q-P/N-rich region spans from aa 857 to 876 (Figure 2.3). This stretch of twenty amino acids contains seven glutamine and seven proline residues, interspersed with small hydrophobic and uncharged amino acids. In humans, ten of these residues are identical, with small changes in position. Several other Q-rich regions are present in the N-terminus of Pum as well and can form aggregates in yeast cells and amyloid fibrils (Macdonald 1992, Salazar, Silverman et al. 2010). Due to the rapid evolution of aggregate-forming glutamine-rich areas of proteins, the overall amino acid compositions tend to be conserved, even if the positions are not. Both human Pum proteins also retain multiple areas rich in glutamine residues, and mouse Pum2 can form aggregates in neurons (Vessey, Vaccani et al. 2006).

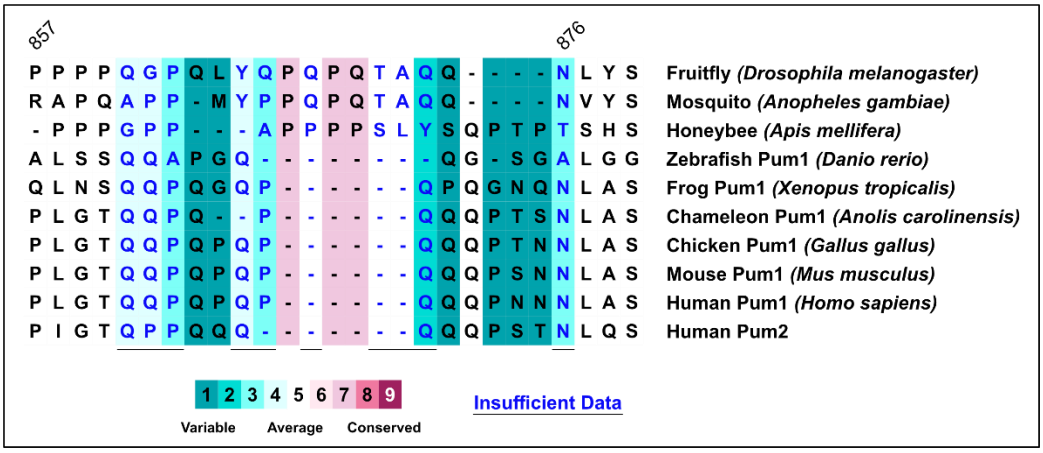


Figure 2.3: The Q-P/N region. A selection of ten different Pum proteins from a ConSurf alignment of 45 different species is shown. Relative conservation for each amino acid position is shaded according to the ConSurf color scheme. Residues with insufficient data to reliably call position are underlined. Amino acid positions for *Drosophila* Pum are listed.

CR1 lies within aa 895 to 913 of *Drosophila* Pum (Figure 2.4). Of these 19 amino acids, 13 are identical in human Pums. Serine 902 is a known phosphorylated residue and is one of the conserved residues (Bodenmiller, Malmstrom et al. 2007). Serine residues make up 32% of CR1, making CR1 relatively serine-rich. This feature could have functional significance, as both human Pums are enriched in serine residues as well.

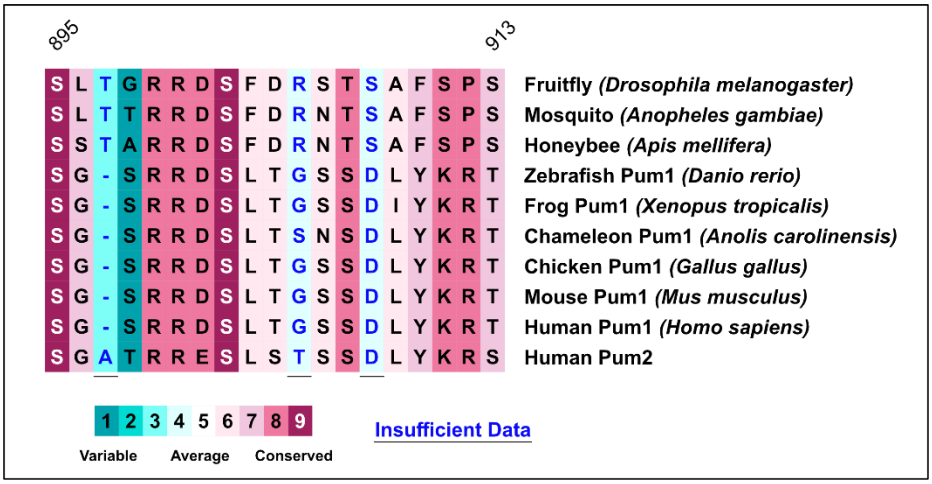


Figure 2.4: Conserved region 1. A selection of ten different Pum proteins from a ConSurf alignment of 45 different species is shown. Relative conservation for each amino acid position is shaded according to the ConSurf color scheme. Residues with insufficient data to reliably call position are underlined. Amino acid positions for *Drosophila* Pum are listed.

CR2 is separated from CR1 by an intervening region comprised of 43% alanine residues. In human Pums, CR1 and CR2 are only about 13 aa apart. In *Drosophila* Pum, CR2 spans from aa 969 to 984 (Figure 2.5). In humans and many other species, an additional nine amino acids are inserted just before I-979. Apart from this divergence, 69% of *Drosophila* CR2 residues are identical to human.

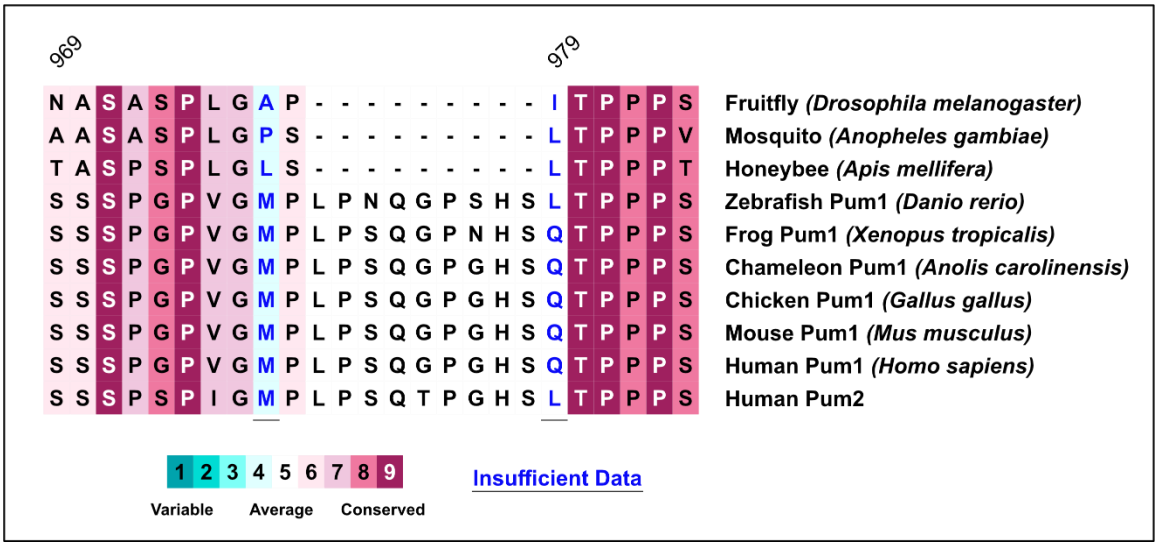


Figure 2.5: Conserved region 2. A selection of ten different Pum proteins from a ConSurf alignment of 45 different species is shown. Relative conservation for each amino acid position is shaded according to the ConSurf color scheme. Residues with insufficient data to reliably call position are underlined. Amino acid positions for *Drosophila* Pum are listed.

CR3, (*Drosophila* Pum aa 992 to 1002), is the most invariant region in all 45 species (Figure 2.6). The initial seven residues, with the exception of R993, are absolutely conserved in both composition and position. For CR3, *Drosophila* appears to be the unusual species in that it harbors the most deviation in the latter half of CR3. Of all aligned species, only *Drosophila* lacks the A/T-K-Y residues. Instead, *Drosophila* Pum uses S-R-Q.

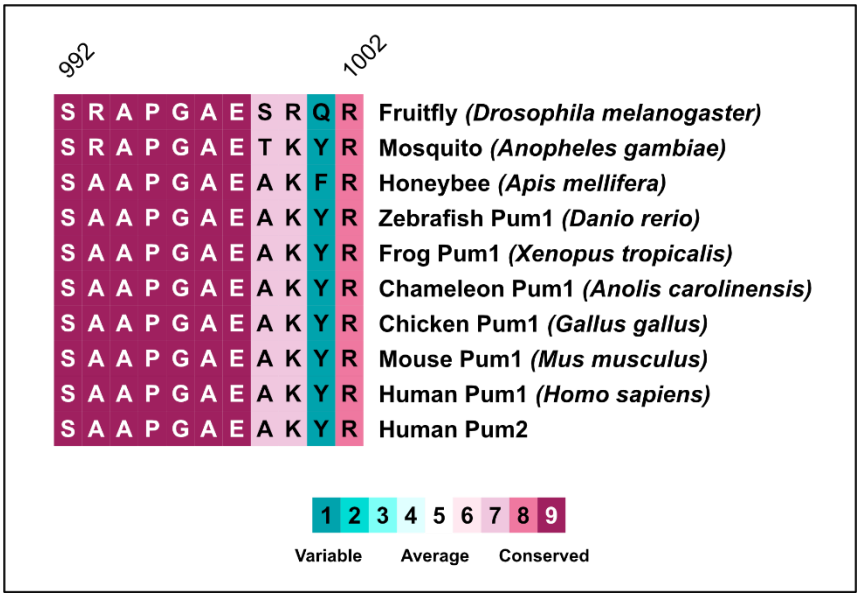


Figure 2.6: Conserved region 3. A selection of ten different Pum proteins from a ConSurf alignment of 45 different species is shown. Relative conservation for each amino acid position is shaded according to the ConSurf color scheme. Amino acid positions for *Drosophila* Pum are listed.

The space between CR3 and CR4 is another region rich in alanine residues, including a stretch of 6 glutamines. In higher organisms, this intervening region is absent, making CR3 and CR4 essentially one longer conserved region.

In flies, CR4 begins at aa1030 and spans 11 residues (Figure 2.7). CR4 is followed by another alanine-rich region that is also absent in most other species, including other insects. In *Drosophila*, CR4 contains several asparagine residues that have evolved into serines in higher organisms. There are two invariant phenylalanine residues as well.

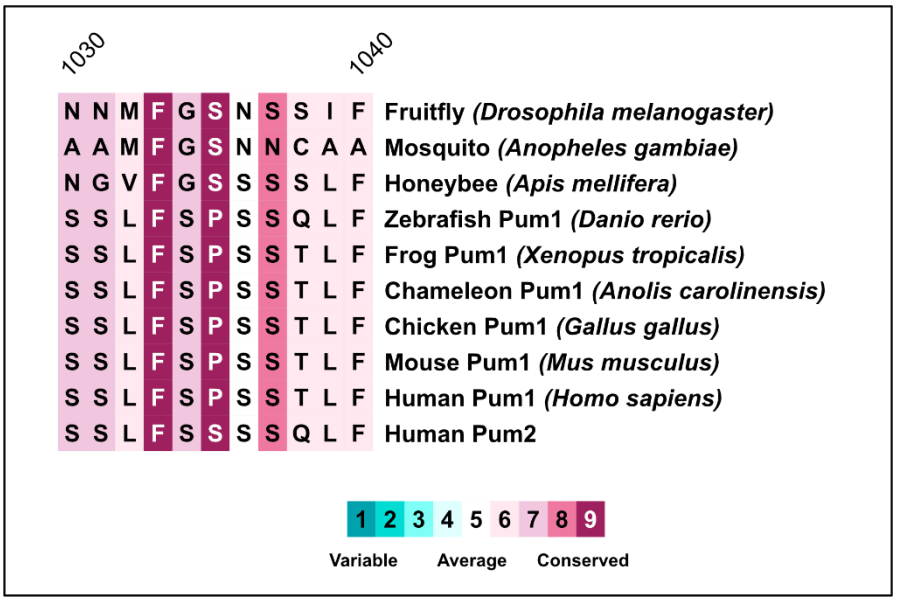


Figure 2.7: Conserved region 4. A selection of ten different Pum proteins from a ConSurf alignment of 45 different species is shown. Relative conservation for each amino acid position is shaded according to the ConSurf color scheme. Amino acid positions for *Drosophila* Pum are listed.

2.4. Functional activity of conserved regions

To test the functional relevance of these conserved regions and identify the area(s) of RD3 responsible for activity, a tethered luciferase assay system was employed. Full-length RD3 or deletion constructs were expressed as fusions to N-terminal bacteriophage MS2 coat protein. MS2 coat protein specifically recognizes and binds as a homodimer to RNA MS2-stem loop structures. A plasmid encoding Nanoluciferase (NLuc) enzyme followed by a set of two MS2 stem loops served as the target substrate. This system is an established and reliable readout for measuring the repressive activity of Pum (Weidmann and Goldstrohm 2012, Arvola, Chang et al. 2020). In this way we were able to direct RD3 constructs to the NLuc reporter target (Figure 2.8). Plasmids were co-transfected with firefly luciferase (FFLuc) as an internal control and fold change was determined as the ratio of NLuc to FFLuc. Repressive activity was determined relative to MS2-eGFP negative control (Figure 2.8 A). Normalizing activity to tethered MS2-eGFP accounts for the natural repressive effect seen when directing any tethered protein to an MS2 stem loop reporter.

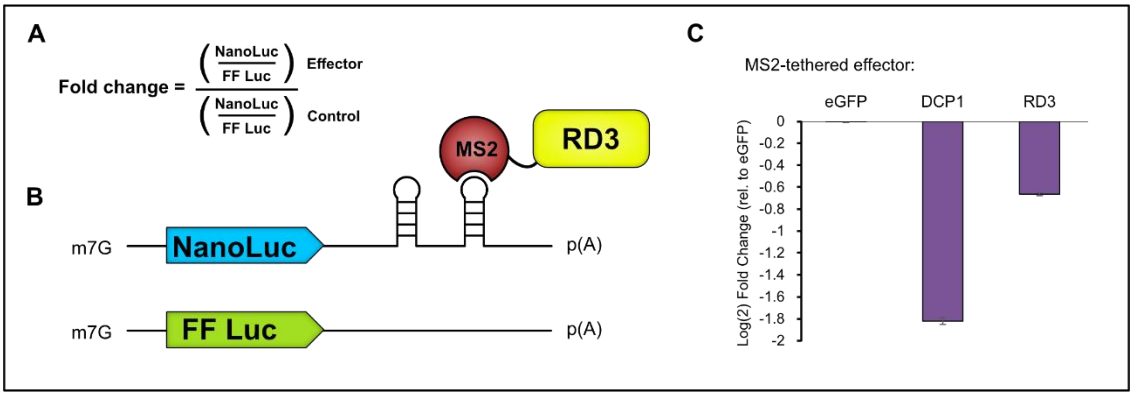


Figure 2.8: The tethered luciferase system. (A) Fold change is calculated by dividing the ratio of Nano luciferase (NLuc) reporter to Firefly luciferase (FF) control for the effector condition to the NLuc:FF ratio for the control condition. FF is used to control for transfection efficiency and general cell health and viability. (B) Schematic representation of the NLuc tethered reporter. The 3' UTR of the NLuc transcript contains 2 MS2 stem loop structures which are recognized by the MS2 viral coat protein. Each RD3 effector is expressed as an MS2 protein fusion, which serves to convey RD3 to the transcript in the absence of an RBD. The FF internal control has a minimal 3'-UTR and is not regulated. (C) Representation of repression. DCP1, a decapping factor, is a positive control for decay of the NLuc reporter. RD3 also causes decay of the reporter. Both effectors are normalized to eGFP. Repression is shown as log(2) fold change.

In this system, WT RD3 repressed the reporter an average log(2) fold change of -0.663 when normalized to eGFP. As a positive control, we included the decapping factor Dcp1, which resulted in an average log(2) fold change of -1.82 when tethered (Figure 2.8 C). After establishing the effectiveness of this system, we used a dual approach consisting of truncations of RD3 and internal deletions of the CRs.

2.4.1. Truncations of RD3

The first constructs tested were simple divisions of RD3, including halves and thirds (Figure 2.9 A). Consideration was taken in ensuring no CR was interrupted in the design of these truncations. The first half consisted of 121 amino acids (847 to 967): the Q-P/N region and CR1, while the second half retained CR2, CR3, and CR4 and comprised 123 amino acids (968 to 1090). The thirds of RD3 were divided into aa 847 to 926 (1/3), aa 927 to 1009 (2/3), and aa 1010 to 1090 (3/3). The 1/3 area of RD3 contained the Q-P/N region and CR1, the 2/3 contained CR2 and CR3, and the last 3/3 contained only CR4. Very little activity remained when RD3 was divided into small regions (Figure 2.9 B). The first half exhibited only about 40% of WT RD3 activity, while

the second half had about 20%. This could have been due in part to a lower level of expression of these constructs when compared to RD3 (Figure 2.9 C).

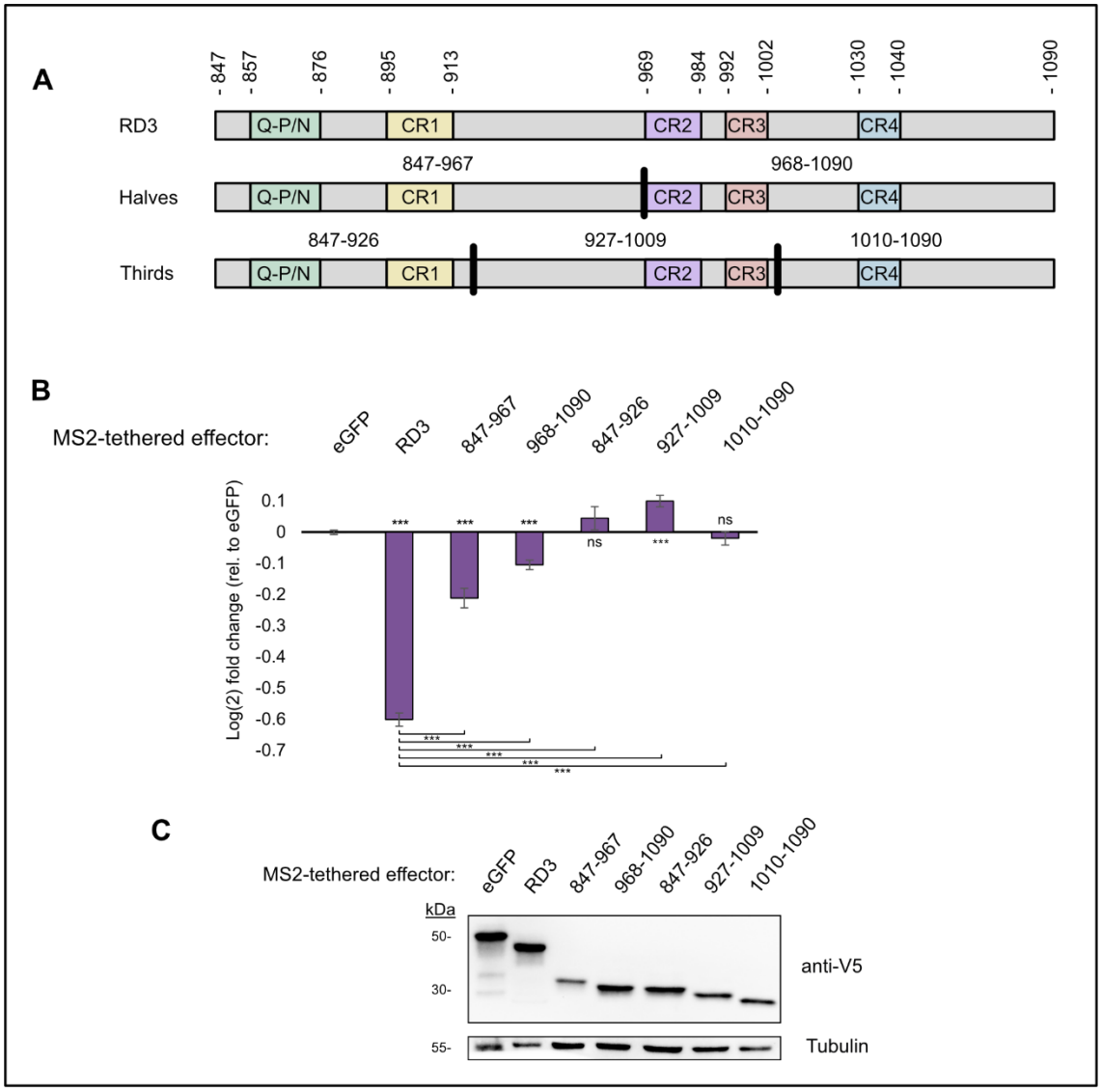


Figure 2.9: Repressive activity of RD3 truncations. (A) Schematic representation of truncations made of RD3. Heavy lines separate where divisions were made for halves and thirds of RD3. Amino acid positions are listed above the diagram which correspond to the numbering of truncations in the graph. (B) Log(2) fold change data showing the mean activity level of RD3 and the truncations tested, with standard error of the mean (SEM) indicated (n ≥ 16). Significance compared to eGFP is shown along the x-axis, while significance compared to WT RD3 is shown at the bottom of the graph. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (C) Western blot showing protein expression levels of all tested constructs. Tubulin is used as a loading control.

A titration of RD3 was done to determine if a slight decrease in expression level could account for the poor activity seen with the truncations (Figure 2.10). RD3 still exhibited nearly full function even at 10 ng transfected DNA with 44% of the protein level compared to RD3 (29% repression at 10 ng, compared to 36% for 100 ng; 81% of WT). At 5 ng transfected only 20% as much protein was expressed and RD3 had 69% activity of WT. The small thirds of the protein, despite having adequate protein expression, had no activity. This indicated that the CRs represented in each truncation were not enough in this context to enact repression.

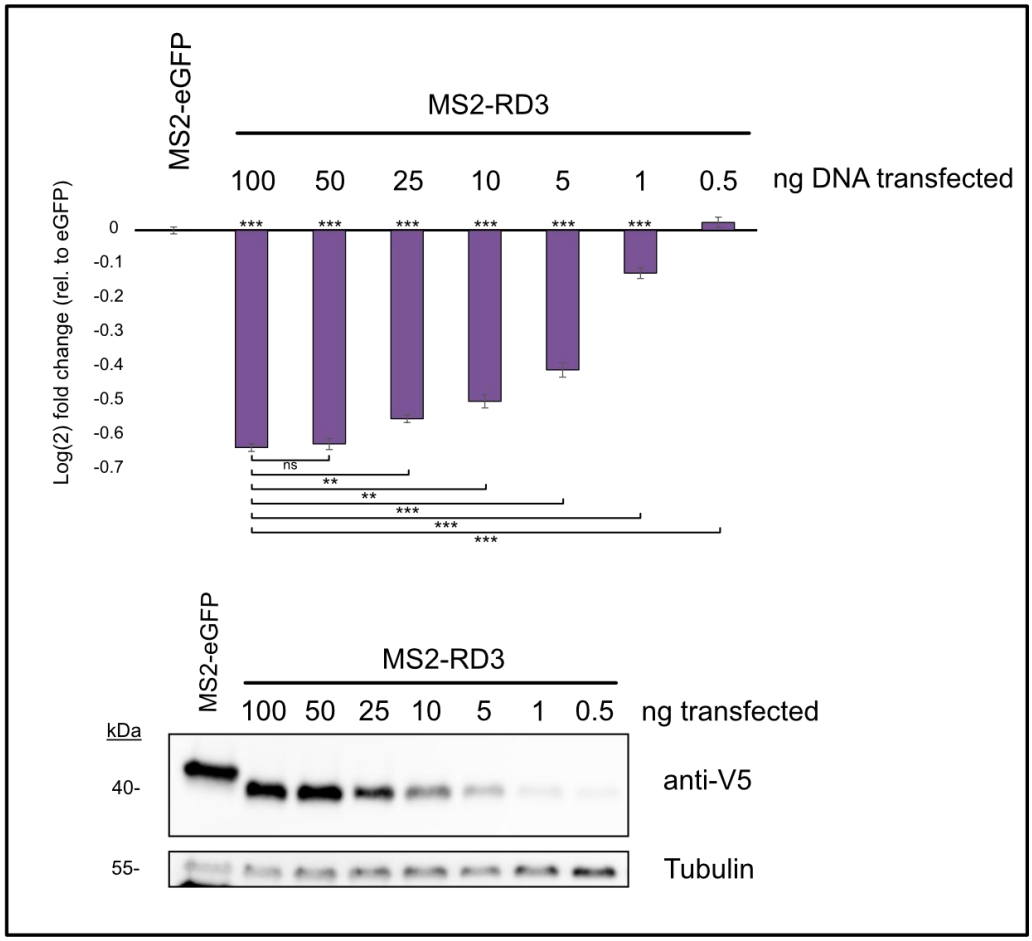


Figure 2.10: Titration of RD3. (Top) Differing amounts of plasmid DNA encoding RD3 were titrated to determine the effective concentration of DNA to protein expression level (n = 4, 4 biological replicates of one experiment). Mean values are plotted with SEM. Significance compared to eGFP is shown along the x-axis, while significance compared to 100ng transfected RD3 is shown at the bottom of the graph. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing expression level of RD3 with tubulin used as a loading control.

Other truncations of RD3 were tested and results are summarized in Figure 2.11. One of the observations made was that the stretch of amino acids between CR4 and the RBD was not necessary for function since the truncation had 100% activity (aa 847 to 1042). The region of RD3 encompassing CR2 and CR3 also did not appear crucial for repression although we did not test removing them without the upstream aa 913 to 969. Curiously, this region of RD3 with no CRs did appear to have an impact on activity when we compared aa 847 to 967 to aa 847 to 926, and aa 927 to 1090 to aa 968 to 1090 (Figure 2.11). There is very little sequence conservation in this area, although there are two serine residues most species possess in varying positions.

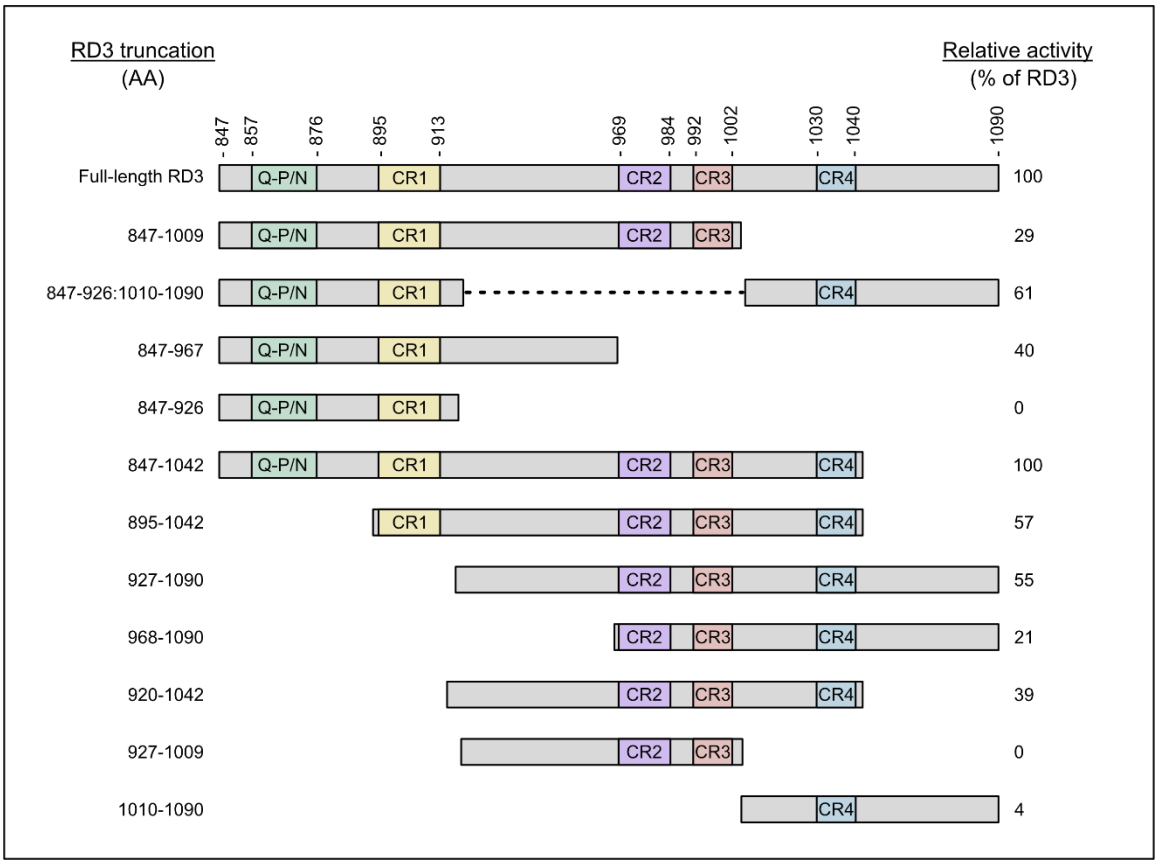


Figure 2.11: Summary of RD3 truncations. Schematic representation of the tested truncations of RD3, with CR positions and amino acid positions shown. The average activity of each construct is shown as a percentage of WT RD3. n ≥ 12 for all constructs.

2.4.2. Deletions of conserved regions

Next, we tested deletions of the conserved regions. Individual deletions of the CRs had varied effects on the repressive activity of Pum RD3 (Figure 2.12). Deleting the Q-P/N region resulted in a relatively small but significant decrease in repression. ΔQ-P/N retained about 84% of wild type RD3 activity. Deleting CR1 had a greater impact on activity- removing this CR resulted in 72% of WT activity. Interestingly, deleting CR2 did not seem to have much of an impact on repression at all. When CR3 was removed, RD3 retained 66% of its wild type activity. The largest effect on activity was seen when CR4 was removed- only 21% of RD3 activity remained. These results indicated that more than one CR might be necessary for RD3's activity. Without structural information, it was not possible to determine if internal deletions were affecting folding or other potential structural motifs RD3 might have. Because RD3 is predicted to be unstructured, we hypothesized the lack of activity was due to the missing amino acids affecting RD3's ability to contact a binding partner needed for repression.

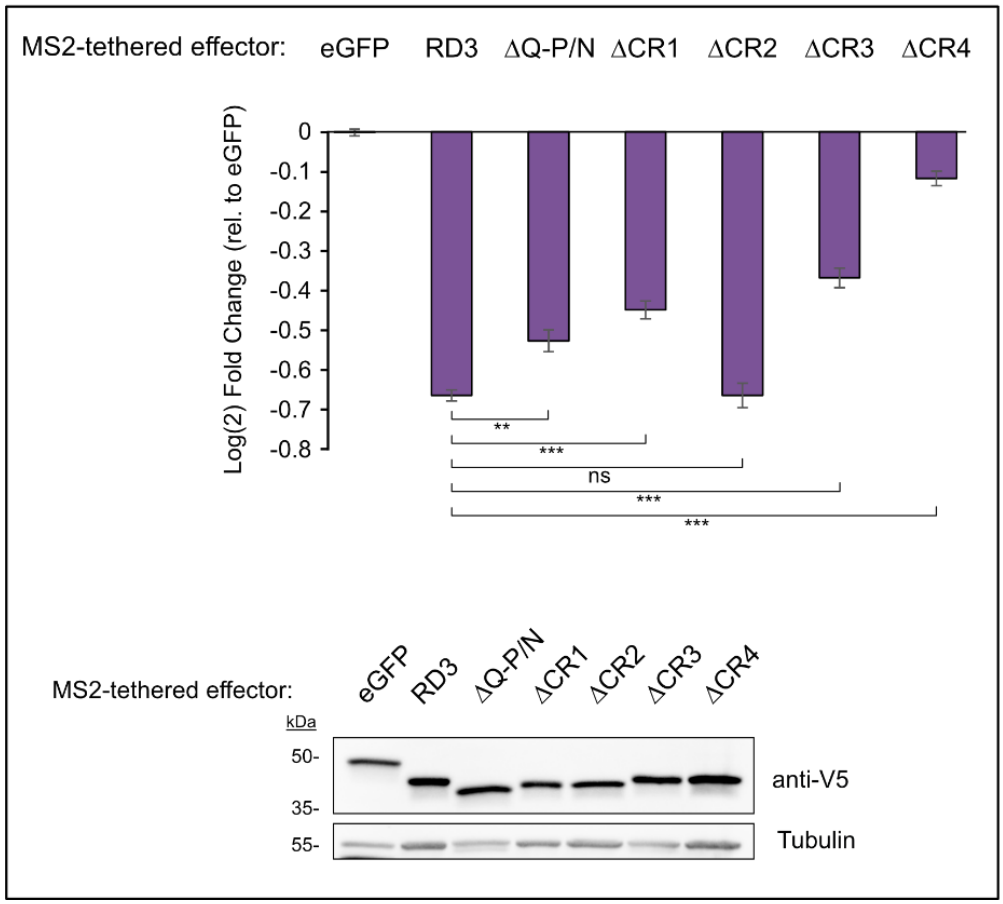


Figure 2.12: Deletions of the conserved regions of RD3. (Top) Graph showing mean relative activity level of RD3 and single truncations of the CRs, with SEM (n ≥ 12). Significance compared to WT RD3 is shown. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing protein expression of all tested constructs, with tubulin used as a loading control.

Since the deletion of any single CR was insufficient to completely abrogate function, we next tested deletions of two or more conserved regions. Deleting the Q-P/N region or CR2 with other CRs did not have an appreciable effect on repression over the single CR deletions, so further analysis was focused on CR1, 3, and 4. However, double deletions involving CR2 had an unexpected effect in that deleting CR2 with another CR mitigated the loss of repression seen with a single CR deletion (Figure 2.13). With the exception of CR4, when any CR was deleted in conjunction with CR2, there was no loss of activity. Removing CR2 with CR4 restored activity over ΔCR4 alone to approximately

50% of WT RD3. However, as our main focus was determining which areas of RD3 contribute to repression, this phenomenon was not explored further.

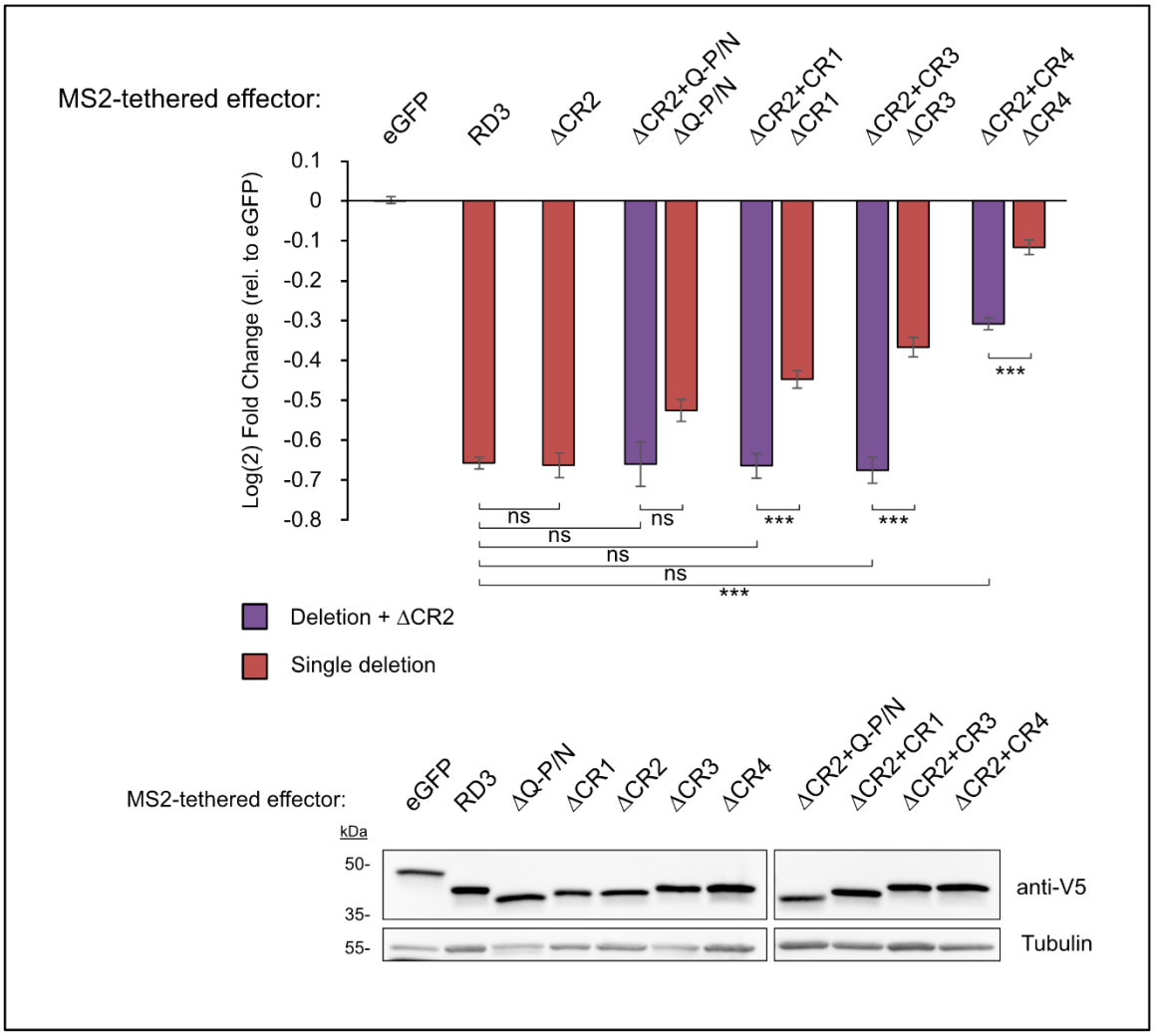


Figure 2.13: Double CR deletions involving CR2. (Top) Mean relative activity of single CR deletions compared to double deletions. Single deletions are labeled and shown in red. Double deletions involving CR2 are labeled and shown in purple (n ≥ 12). Error bars represent SEM. Significance between pairs of comparisons is shown. Significance compared to WT RD3 is shown at the bottom of the graph. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing protein expression of all tested constructs, with tubulin used as a loading control.

Since CR1, CR3, and CR4 appeared to contribute the most to RD3's repressive activity, we focused on double deletions involving these CRs. Deletions of CR4 with CR1 were highly effective in alleviating repression by RD3, and reduced activity to

approximately 4 percent of wild type RD3 (Figure 2.14). Although ΔCR3+CR4 was also striking, it was not statistically different than deleting CR4 alone. In addition, deleting CR1 in conjunction with CR3 did not have an additive effect on repression. This indicated that CR3 was not as important as CR1 and CR4 for RD3's repressive function. We hypothesized that CR4 could be a main point of contact with a protein partner and that CR1 might also contact a partner in a secondary fashion or strengthen CR4's association with it.

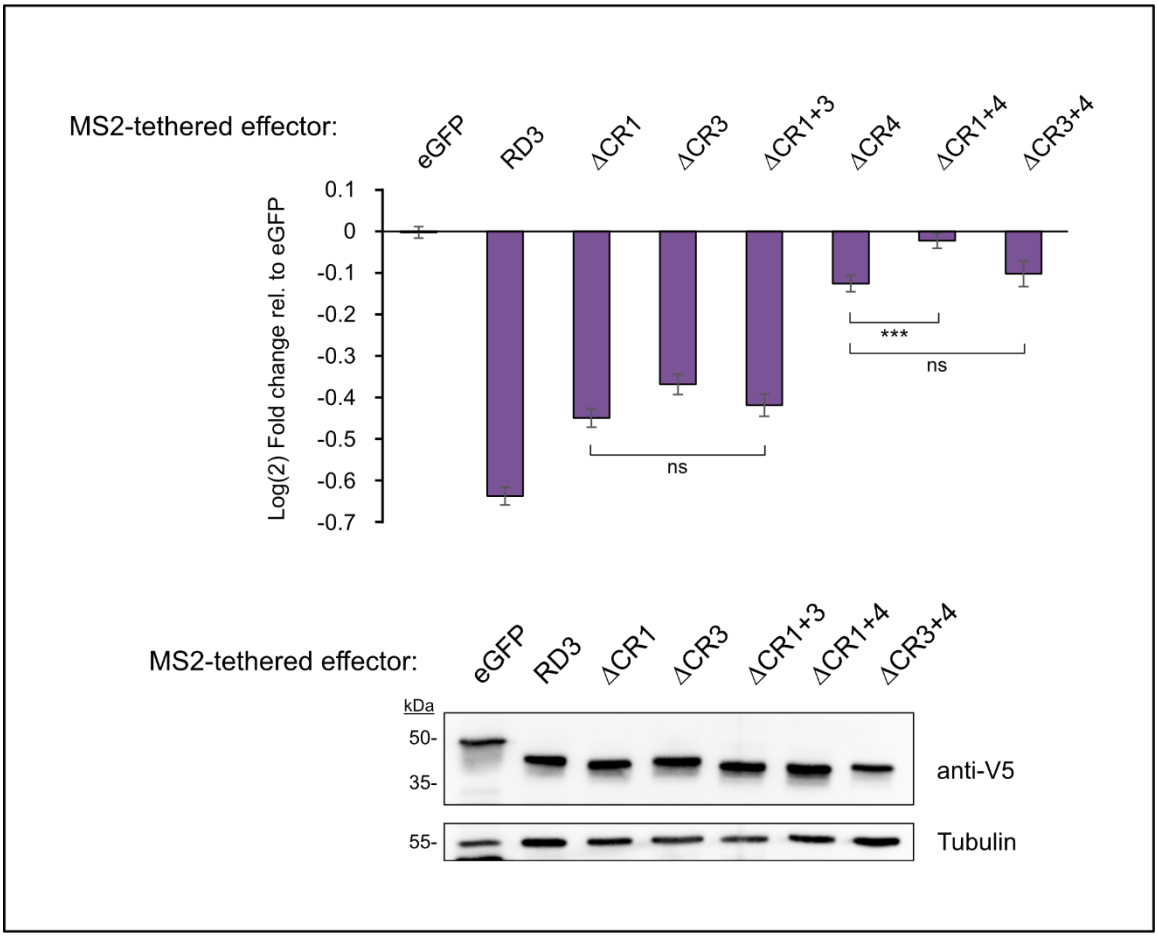


Figure 2.14: The effects of double CR deletions on RD3 activity. (Top) Graph showing mean relative activity of deletions involving 2 CRs of RD3 (n ≥ 12). Error bars represent SEM. Significance compared to ΔCR1 and ΔCR4 is shown. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing protein expression of all tested constructs, with tubulin used as a loading control.

2.4.3. Alanine scanning of conserved regions

To identify the important residues within CR1 and CR4, an alanine scanning strategy was employed. We reasoned that if a group of mutations were to cause an equal loss of activity as deleting the entire CR, it would indicate that one or more of the amino acids in the group could be a site of protein interaction. Since CR1 is 19 amino acids in length, groups of 3 to 4 residues were mutated to alanine (Figure 2.15 A), while groups of 2 to 3 residues were used for CR4 (Figure 2.16 A).

When alanine mutations in CR1 were tested, the results were similar for all groups. Every group reduced repressive activity, and none of the mutants were significantly different from ΔCR1. Three groups showed statistical difference from WT RD3: RRD>AAA, SAF>AAA, and SPS>AAA. These results indicated that any of the tested residues could be important.

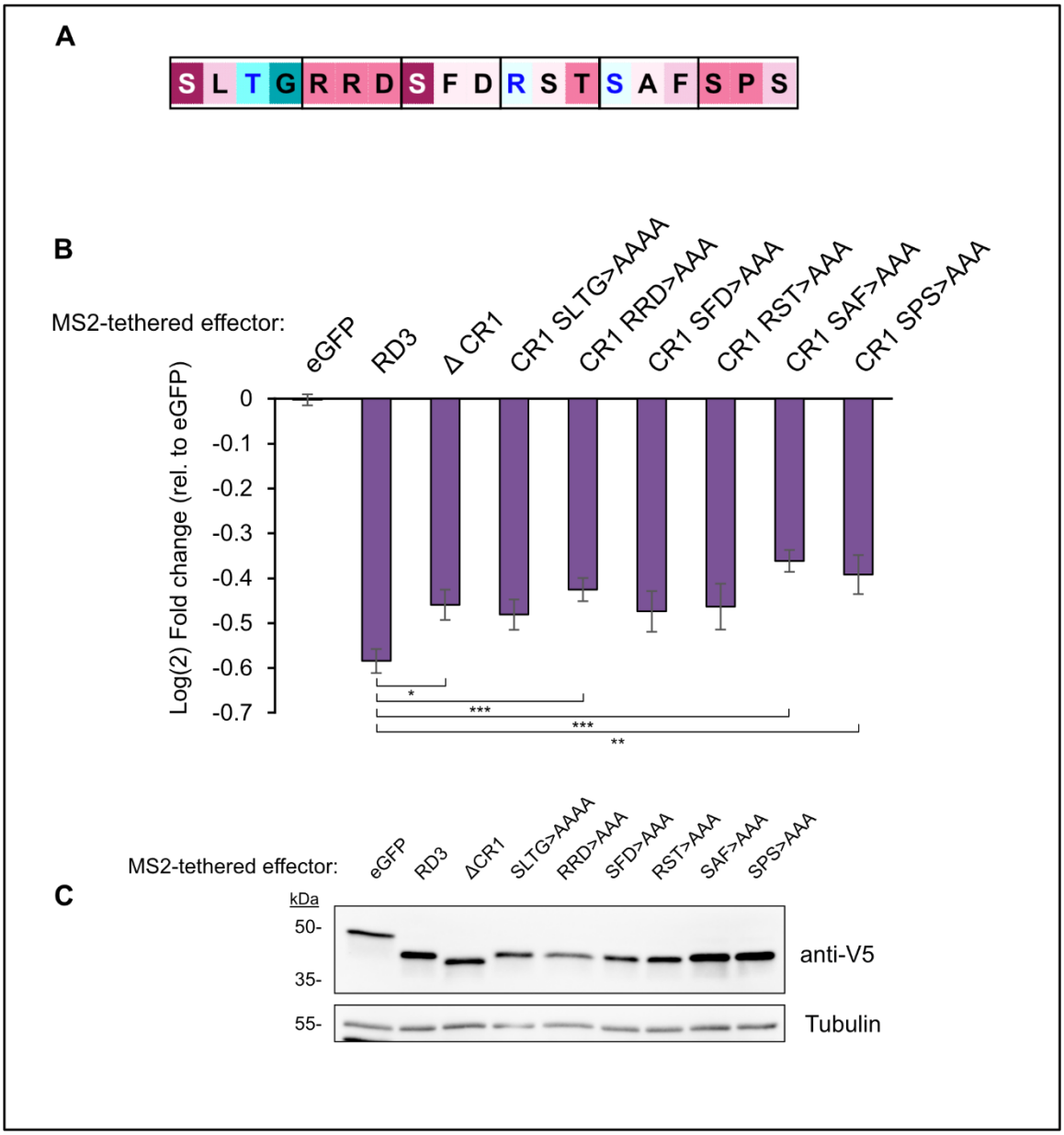


Figure 2.15: Alanine scanning of CR1. (A) Diagram of groups of alanine mutations. The sets of mutations tested are indicated by boxes. (B) Mean relative activity level of CR1 alanine mutations is shown compared to WT RD3 and ΔCR1 (n ≥ 11). Error bars represent SEM. Significance compared to WT RD3 is shown. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (C) Western blot showing protein expression of all tested constructs, with tubulin used as a loading control.

Five groups of alanine mutations in CR4 were tested (Figure 2.16). Interestingly, the first set of two asparagine residues caused an increase in activity when mutated to alanine, increasing repression from 31%, when compared to eGFP, to 43%. In most

species, these two amino acid positions are serine residues (Figure 2.7). The sets GS>AA and NS>AA had no effect on activity. The last two sets, MF>AA and SIF>AA both showed a substantial loss of repression compared to WT RD3. To test whether the loss of activity was due to mutating the serine residue, and to avoid the potential for neighboring serines to “compensate” for its loss, all three serines in CR4 were mutated to alanine. This resulted in a slight increase in activity when compared to the group mutation of SIF>AAA (Figure 2.16 B). This implied that the serine residue was not the most important factor.

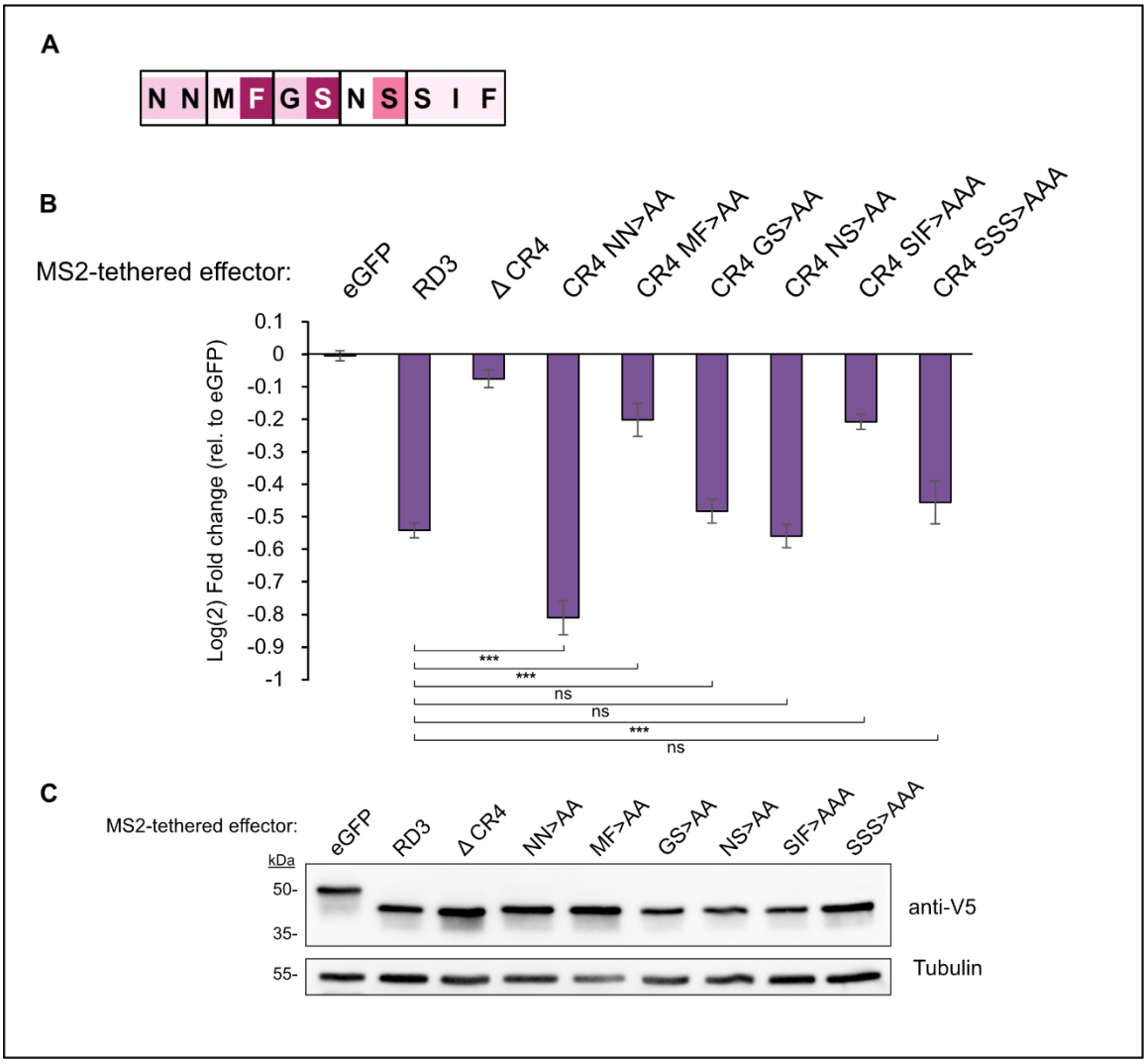


Figure 2.16: The effects of CR4 alanine mutations on RD3 activity. (A) Diagram of groups of alanine mutations. The sets of mutations tested are indicated by boxes. (B) Mean relative activity level of CR4 alanine mutations is shown compared to WT RD3 and ΔCR4 (n ≥ 12). Error bars represent SEM. Significance compared to WT RD3 is shown. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (C) Western blot showing protein expression of all tested constructs, with tubulin as a loading control.

Since an isoleucine to alanine change is a conservative mutation, as is methionine to alanine, we reasoned that the two phenylalanine residues present in both groups were important. Additionally, phenylalanine residues are often seen at protein-protein interfaces due to their bulky and hydrophobic nature. Mutating each

phenylalanine residue separately resulted in a loss of approximately half of RD3's activity. When both F1033 and F1040 are mutated together, RD3 activity was destroyed (Figure 2.17).

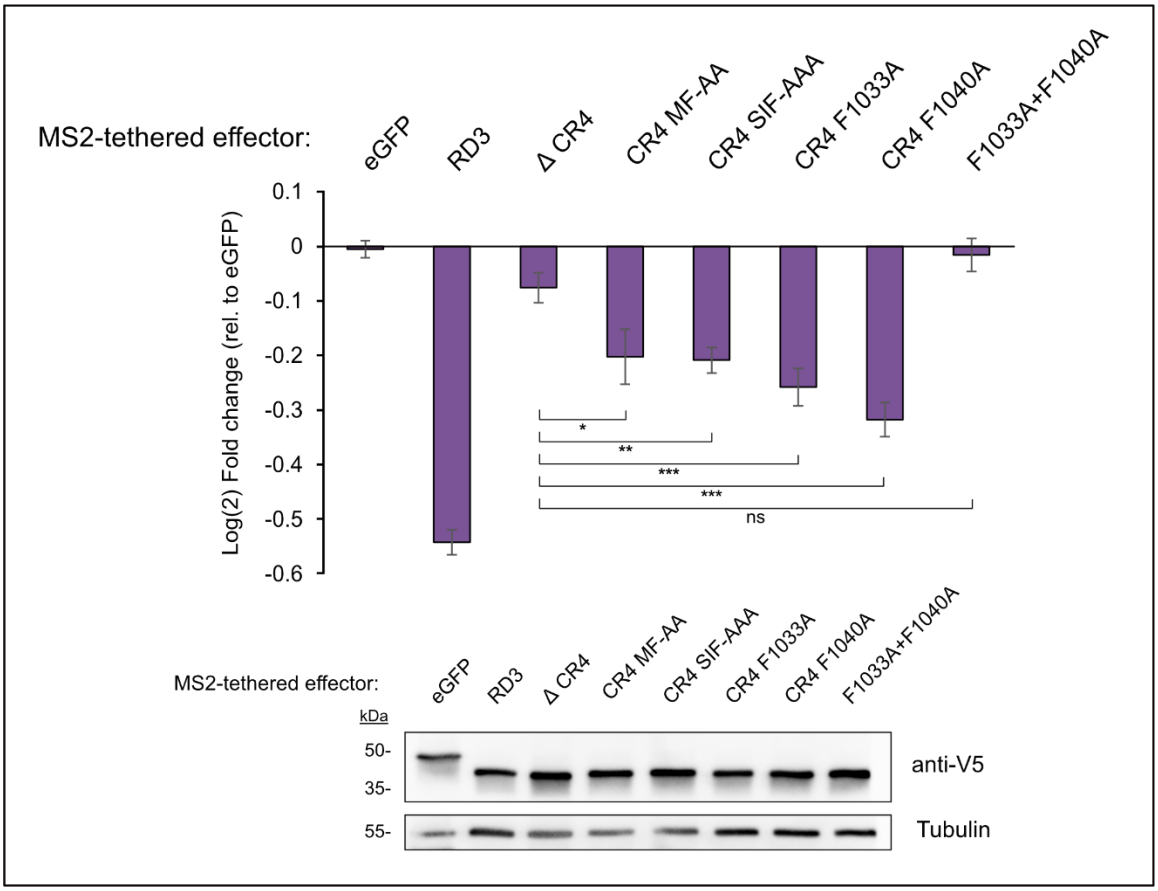


Figure 2.17: Mutations of the phenylalanine residues in CR4. (Top) Mean relative activity level of phenylalanine mutations compared to the first mutation sets tested, with SEM ($n \geq 12$). Significance compared to ΔCR4 is shown. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing protein expression of all tested constructs, with tubulin used as a loading control.

The effects of deleting CR1 alone were minimal to RD3's repressive activity, yet when CR4 was deleted in combination with CR1, repression was lost entirely. We sought to identify the residues in CR1 which appeared to work in conjunction with CR4. Since mutating both phenylalanine residues in CR4 eliminated repression, we focused on mutations involving F1033 with CR1 groups where we could observe any effect of synergism. Each single CR1 mutation was compared to the combined mutations with

F1033A (Figure 2.18). The expected combined additive effect of the loss of functional residues with RRD>AAA and F1033A would be about 33% activity of WT RD3. With 20.5% activity remaining, only the RRD>AAA + F1033A group appeared to exhibit synergism. The RRD>AAA group was also an important group when mutated in WT RD3 (Figure 2.15 above for CR1).

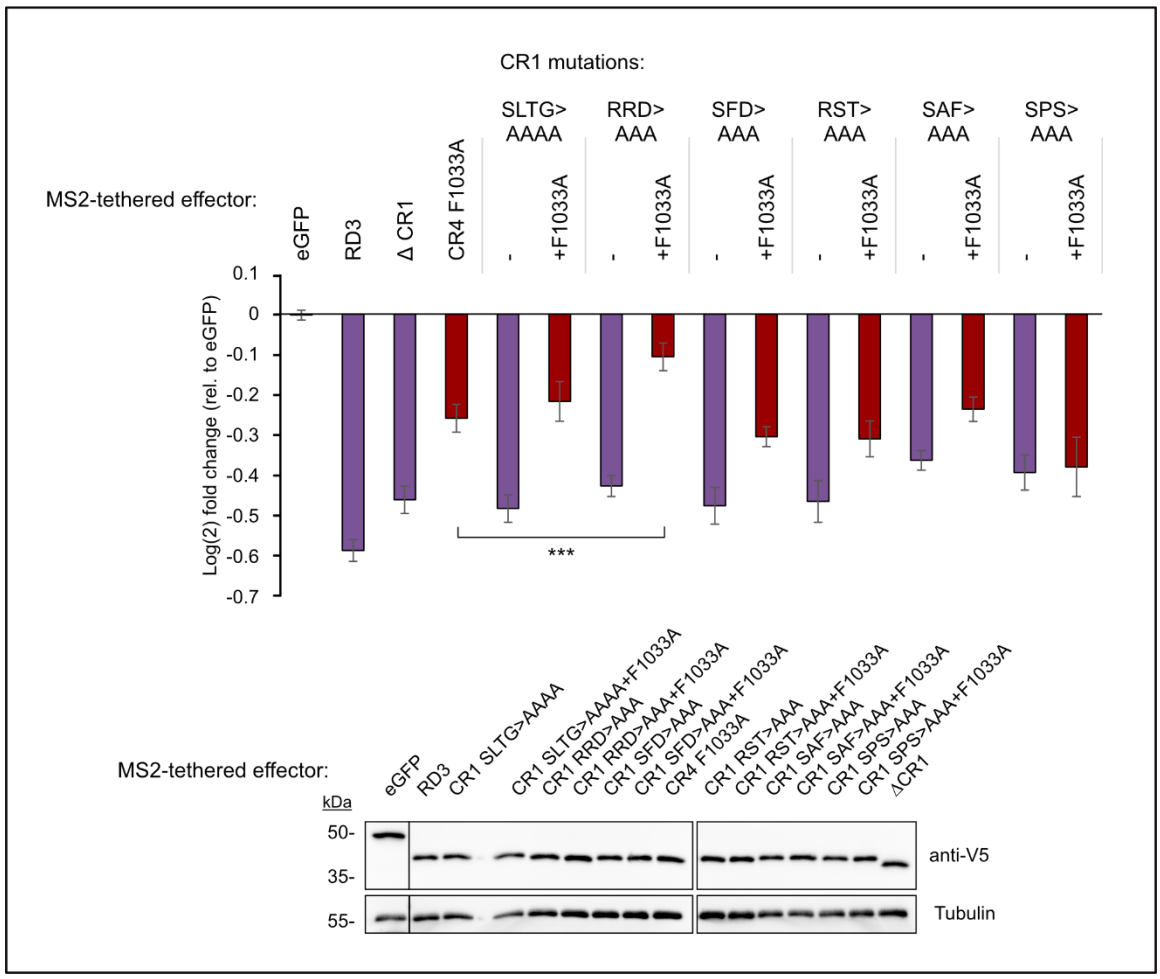


Figure 2.18: Effects of alanine mutations involving CR1 and CR4 F1033A. (Top) Mean relative activity of CR1 alanine mutations either alone (purple) or in combination with CR4 F1033A (red). n ≥ 11. Error bars represent SEM. Significance compared to the single mutation of CR4 F1033A is shown. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing protein expression of all tested constructs, with tubulin used as a loading control.

2.4.4. Sufficiency of CR1 and CR4

Having established that CR1 and CR4 are necessary for RD3's activity, we next asked if they were sufficient. Multimers of CR1 and CR4 were cloned as MS2-eGFP fusions with the eGFP being added to improve protein expression levels. Each set of CR1+4 was separated by a flexible serine-glycine linker 28 amino acids in length, with the same linker between CR1 and CR4 (Figure 2.19).

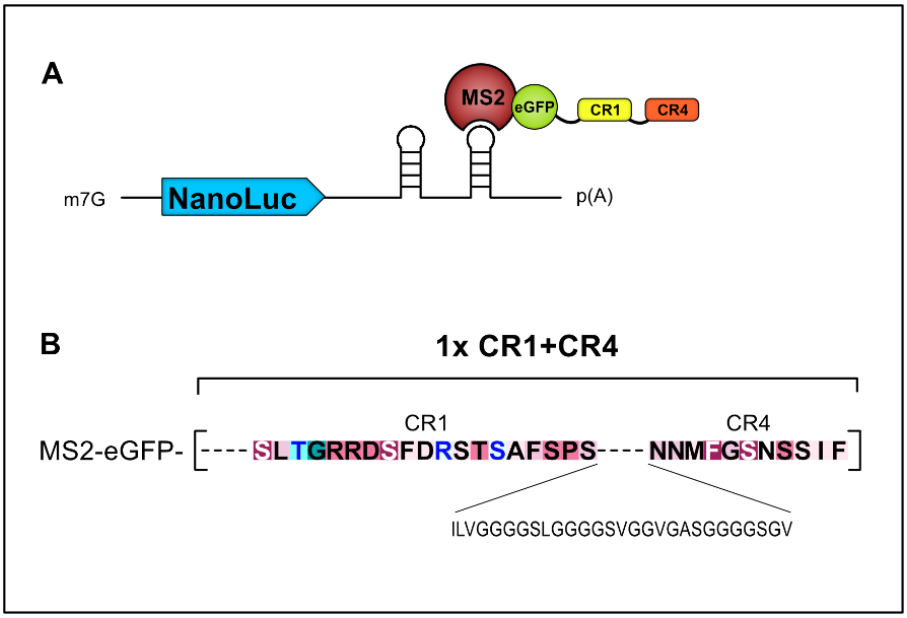


Figure 2.19: The CR1+4 constructs. (A) Schematic representation of the luciferase assay reporter with MS2 stem loops, and the MS2-eGFP CR1+4 multimer constructs. (B) Amino acid sequence identity of the multimer constructs. Following the MS2-eGFP sequence is a flexible 28 amino acid linker preceding CR1, represented by dashes. The same linker separates CR1 and CR4. Brackets denote the repeating unit considered one unit of the multimer. The identity of the linker sequence is shown below CR1 and CR4.

Multimers of CR1+4 are sufficient to repress a reporter mRNA (Figure 2.20). A small amount of activity was seen with a single set of CR1+4, however the construct was not expressed well. As the number of CR1+4 repeats was increased, repression also increased, up to a level that surpassed WT RD3 activity, with 5x CR1+4. 6x CR1+4 also had more repressive activity than WT RD3, but it was not expressed as well as the 5x construct.

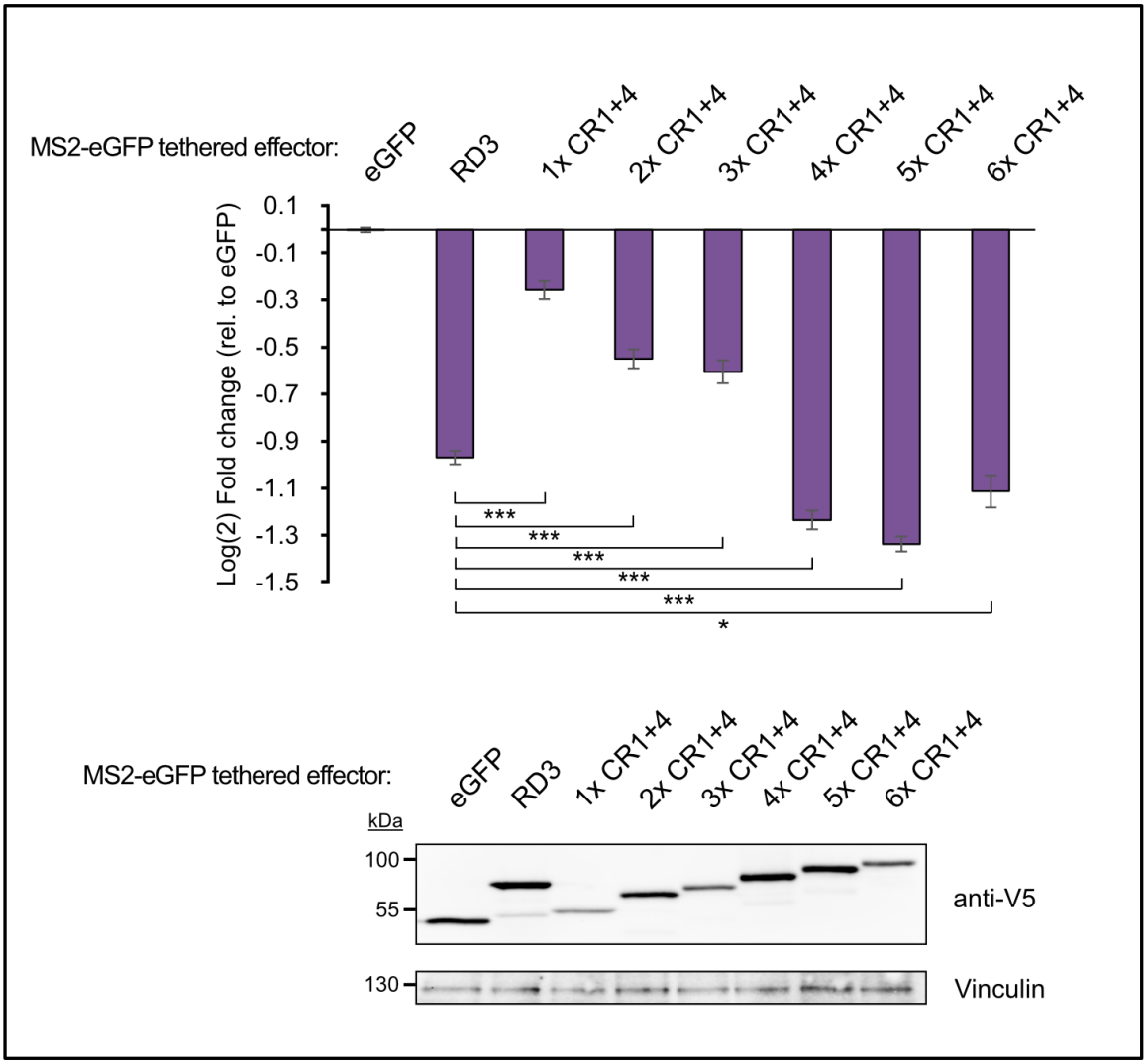


Figure 2.20: Repressive activity of the multimers of CR1 and CR4. (Top) Mean relative activity level of CR1+4 multimers compared to WT RD3 (n ≥ 12). Error bars represent SEM. Significance compared to WT RD3 is shown. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing protein expression of all tested constructs, with vinculin used as a loading control.

To demonstrate the importance of the phenylalanine residues in these constructs, a 5x CR1+4 mutant was made in which both phenylalanine residues were mutated to alanine (5x CR1+4 FF mut). Although the mutant constructs retained a significant amount of repressive activity, it was due in part to their higher expression levels. To determine how much activity could be attributed to the higher expression level of the mutant, we used western blots normalized to tubulin to determine relative expression levels between the 5x CR1+4 and its mutant. The relative expression levels

from three independent titrations were averaged to give an overall equivalency estimate. By this calculation, 100 ng of the 5x CR1+4 construct expressed at the same level as 24 ng of the 5x CR1+4 FF mutant. At an equivalent level of protein expression, the 5x FF mutant had very little activity (Figure 2.21). This further established the importance of the phenylalanine residues in CR4.

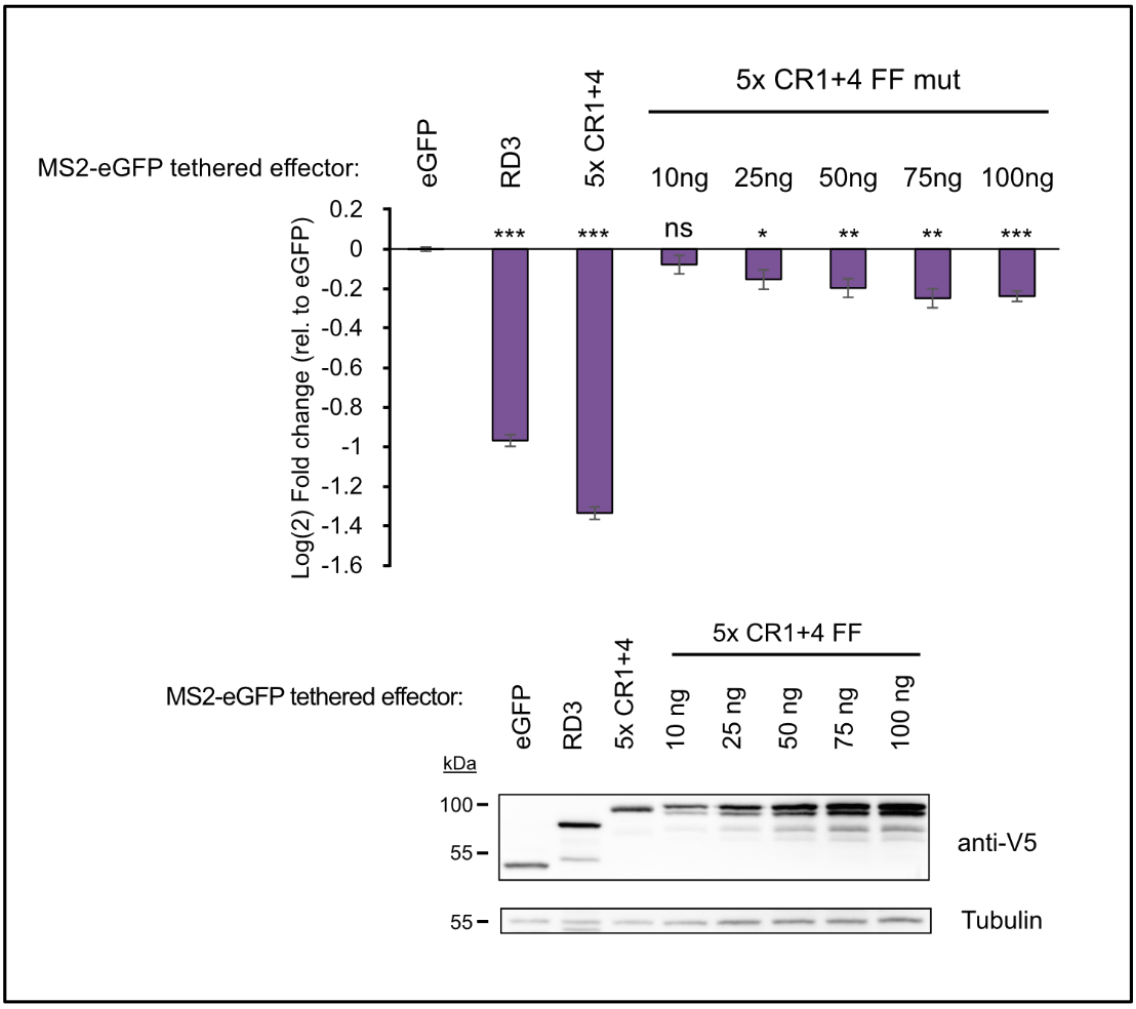


Figure 2.21: Titration of the 5x CR1+4 FF mutant. (Top) Mean relative activity level of increasing amounts of transfected 5x CR1+4 FF mut DNA, compared to WT RD3 and 5x CR1+4 ($n \geq 12$). Error bars represent SEM. Significance compared to eGFP is shown along the x axis. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (Bottom) Western blot showing protein expression of all tested constructs, with tubulin used as a loading control.

2.5. Summary and discussion

In this study, we found that RD3 has several conserved regions. The most important areas for RD3 repressive function were CR1 and CR4, specifically two phenylalanine residues within CR4, F1033 and F1040. Using a tethered reporter system, we were able to demonstrate that these two phenylalanine residues are necessary for RD3's activity. We were also able to show that CR1 and CR4 were sufficient for activity. Deleting CR1 alone had a small effect on activity and reduced RD3 function to 65% of WT RD3. Removing both CR1 and CR4 eliminated RD3 activity. Activity was also lost when CR4 was deleted, or when both phenylalanine residues were mutated.

Curiously, mutating both phenylalanines in CR4 had a more detrimental effect than deleting CR4 entirely. If the position of CR4 is important to its function, it is possible that the deletion of CR4 brings other residues into proximity that may fractionally compensate for its loss. The N-terminal side of CR4 is alanine rich with several glutamines, and this stretch of residues was absent in the other 41 species we examined. The C-terminal side of CR4 is not conserved in amino acid position, but most species contain serine, proline, and tyrosine or threonine interspersed with small hydrophobic residues. A truncation of RD3 which removed all the amino acids following CR4 retained full activity (aa 847 to 1042, Figure 2.11), although this construct was not tested with a CR4 deletion.

Equally intriguing is the observation that deleting CR2 in combination with other CRs appeared to restore activity over the single deletion alone. This implies that CR2 has an auto-regulatory function. There is precedence for this behavior in full-length Pum where PCMa and PCMb exhibit regulatory activity (Weidmann and Goldstrohm 2012). Weidmann and Goldstrohm showed that expressing a construct of PCMb-RD3 alleviates repression by RD3. However, the mechanism of auto regulation remains unexplored. A simpler explanation could be that removing CR2 allows RD3 to adopt a conformation that improves contacts with a binding partner. Alternatively, CR2 might contact another regulatory protein. If this were the case, however, one would expect that removing only CR2 would improve RD3's repressive activity over that of WT, which the data does not show. It also does not appear that RD3's activity is saturated since the NN>AA mutation in CR4 enhances activity to 137% of WT. In the case of the NN>AA mutation, it is possible that removing the larger asparagine residues could allow for the nearby

phenylalanines to better contact the co-repressor protein(s). Evolution may have already accomplished this, as higher species have serine residues in these positions. If CR2 is exerting auto-regulatory behavior, it could help to explain why the 5x CR1+4 construct surpassed WT RD3 activity. With no regulatory domain, 5x CR1+4 may be able to function at their full potential.

In the native RD3, CR1 and CR4 are separated by 135 amino acids. In the 5x CR1+4 construct, they are 28 residues apart. Even though 2x CR1+4 was adequately expressed, it has only 65% of WT RD3 activity. It was only when 4 or more groups of CR1+4 were expressed that repression reached the level of WT RD3. This could indicate that the spacing between the CRs is important. It is also possible that the other CRs that contribute to repressive activity in the context of WT RD3 are needed for full function in a minimal construct such as CR1+4. Indeed, both the Q-P/N region and CR3 had significant effects on WT RD3 when they were deleted, although they were not essential.

The observation that RD3's repressive ability did not depend upon CR3 was curious given the high level of conservation seen between species in CR3. *Drosophila* is the only species we examined using ConSurf with an extensive stretch of alanine and other residues separating CR3 and CR4. It may be that the importance of CR3 evolved alongside the elimination of this stretch of amino acids. Alternatively, perhaps the unique amino acids S-R-Q in *Drosophila* CR3 diminished its functionality in fly and other species escaped this mutation.

Although the Q-P/N region is not as relevant for RD3's activity, intrinsically disordered regions (IDRs) such as this can have roles in oligomerization, granule formation, or function as flexible spacer regions (van der Lee, Buljan et al. 2014). One study used similar regions present in RD1 and RD2 to show that these Q-rich domains along with the N-terminus of Pum can form aggregates in yeast (Salazar, Silverman et al. 2010). Pum has also been identified as a component in P-bodies, the formation of which are driven by protein aggregation-prone domains and RNA interactions (Barbee, Estes et al. 2006). It is possible that the Q-P/N region in RD3 functions in a similar manner and that the decrease in activity seen when it was deleted was caused by the lack of an additional Pum molecule to help enact decay.

Additionally, it is not known whether the mutations or deletions are affecting the localization or storage or overall folding of the protein. RD3 is predicted to be disordered but it may adopt a structure upon interaction with a protein partner. Such induced folding is common with unstructured and intrinsically disordered proteins (Uversky 2019).

Another, perhaps simpler explanation for the dispensable nature of individual conserved regions is that each one may be only partially responsible for a larger interface of multiple contacts with protein partners. This is a common theme for proteins interacting with large complexes, and is seen with human Pum proteins in their interactions with the CNOT complex as well (Enwerem, Elrod et al. 2021).

2.6. Materials and methods

Cell culture

Drosophila Dmel2 cells (Invitrogen) were cultured in Gibco SF-900III media with no additional supplements. Cells were passaged every 3 to 4 days and discarded once they reached passage 30. *Drosophila* Dl1 cells were a gift from Eric Wagner. Dl1 cells were cultured in Schneider's *Drosophila* media (SDM, Gibco) containing additional glutamine (1x GlutaMAX, Gibco), 1x anti-microbial/anti-fungal (Anti/anti, Thermo Fisher), and 10% heat inactivated fetal bovine serum (FBS, GenClone). FBS was heat inactivated by incubating at 56 to 60°C for 30 minutes. DL1 cells were passaged every 3 to 4 days and discarded after passage number 40. All *Drosophila* cell lines were cultured at 25°C in a humidified incubator. DL1 cells were frozen from 80% density in 10-cm dishes, collected and centrifuged in 15 mL tubes at 800 x g for 3 minutes. 2 mL of used media was reserved; 500 µL DMSO and 2.5 mL of complete SDM were added to the tube and cells were resuspended. Cells were aliquoted into 3 cryogenic storage tubes, placed in a Styrofoam tray and stored at -80°C for 2 to 7 days before being transferred to liquid nitrogen.

Luciferase assays

100 µL of *Drosophila* cells were seeded in 96-well plates at a density of 1.5 x 10^6 per mL. Four wells in each condition served as biological replicates for luciferase measurements, while three wells in each condition were reserved for protein analysis. Transfection master mixes were set up using 1 ng of Firefly Luciferase (FF) per well (5 ng for Dmel2), 1 ng of Nano Luciferase (NLuc) reporter (5 ng for Dmel2), and 98 ng of either effector plasmid or balanced with empty pIZ plasmid (90 ng for Dmel2). FuGene HD transfection reagent (Promega) was used at a 4:1 ratio (0.4 µL/well) and serum free media was added to master mixes to bring the level of master mix transfected per well to 10 µL. Master mixes were incubated for 5 to 10 minutes before adding to wells of plated cells. Plates were incubated for 48 to 72 hours before harvest and analysis. For analysis, 100 µL of OneGlo Firefly reagent (Promega) was added to each well, mixed and incubated at room temperature for 10 minutes. Luminescence was measured on a Glo-Max instrument under default settings. 100 µL of NanoGlo Stop-n-Glo reagent

(Promega) was then added to each well, mixed and incubated at room temperature for 15 minutes before reading luminescence.

Data analysis

Luciferase assays: Samples were normalized for potential variation in transfection efficiency by calculating the ratio of NLuc to FF. The repression activity of each effector was calculated as fold change by dividing each sample well NLuc:FF ratio by the average NLuc:FF ratio of the control condition (eGFP, R7 RNA binding mutant, or ACA mutant reporter). Repression activity was then graphed as log(2) fold change. Statistics for luciferase assays were performed using Student's unpaired T test with either equal or unequal variance as dictated by F-test results. Significance calling was based on the convention $p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***. Each experiment consisted of 3 to 4 biological replicates and each experiment was repeated at least 3 individual times.

Westerns

Three wells of each condition were harvested by pipette-mixing and transferring cells to 1.5mL Eppendorf tubes. Cells were centrifuged at 900 x g for 4 minutes, media was removed, and cells were resuspended in RIPA buffer (25 mM Tris pH 7.6, 1 mM EDTA, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) with additional 2x protease inhibitors (Roche). Lysate was collected by centrifuging at 21,000 x g for 10 minutes and removing the supernatant. Protein concentration was measured using the Bio-Rad DC protein assay kit according to manufacturer's directions.

10 µg of total protein extract was incubated with an equal volume of 2x SDS loading buffer at 85°C for ten minutes. Samples were loaded on 12% SDS-PAGE gels (either laboratory-made or purchased precasted from BioRad) and electrophoresed at 120 to180 volts until the dye front reached the bottom of the gel. Gels were transferred to PVDF membranes (Millipore) for either 1.5 hours at 65 volts or overnight at 35 volts. Following transfer, membranes were either blocked in blotto (1x PBS, 0.1% Tween, 5% dried milk) for 1 hour or allowed to dry completely, then primary antibodies were applied for one hour at room temperature, or incubated overnight at 4°C. Membranes were washed three times for 10 minutes in blotto, then secondary antibodies were applied for 1 hour at room temperature. Membranes were washed three times for ten minutes each

in blotto, then rinsed with 1x PBS before applying chemiluminescent substrate (Pierce-Thermo Fisher, or Immobilon- Millipore) Blots were imaged using the auto exposure setting on a ChemiDoc touch (BioRad).

Plasmids

Archive number	Plasmid name	Creator
CAW061	pIZ MS2CP-GFP V5H6	Chase Weidmann
RMA0053	pIZ MS2-Dcp1 V5H6	Rene Arvola
CAW314	pIZ MS2 Pum RD3i V5H6	Chase Weidmann
CAW313	pIZ MS2 Pum RD3ii V5H6	Chase Weidmann
RMA0008	pIZ Dm Pum RD3-RBD	Rene Arvola
pBA017	pAc5.4 NLuc 2xMS2 BS	Brian Alzua
CAW023	pAc5.1 FFLuc2 minimal 3'-UTR	Chase Weidmann
RJH63	pIZ MS2 RD3 ▲FA2	Rebecca Haugen
RJH64	pIZ MS2 RD3 ▲Q+FA2	Rebecca Haugen
RMA0001	pIZ MS2-RD3 -QN	Rene Arvola
RJH66	pIZ MS2 RD3 ▲FA1+CR3	Rebecca Haugen
RJH67	pIZ MS2 RD3 ▲FA2+CR3	Rebecca Haugen
RJH68	pIZ MS2 RD3 ▲FA4+CR3	Rebecca Haugen
RJH72	pIZ MS2 RD3 ▲1/3	Rebecca Haugen
RJH73	pIZ MS2 RD3 ▲2/3	Rebecca Haugen
RJH77	pIZ MS2 RD3 ▲FA1	Rebecca Haugen
RJH78	pIZ MS2 RD3 ▲FA4	Rebecca Haugen
RJH80	pIZ MS2 RD3 1/3 alone, 1-79	Rebecca Haugen
RJH81	pIZ MS2 RD3 3/3 alone, 163-243	Rebecca Haugen
RJH82	pIZ MS2 RD3 ▲3/3	Rebecca Haugen
RJH87	pIZ MS2 RD3 ▲FA2+1	Rebecca Haugen
RJH88	pIZ MS2 RD3 ▲FA2+4	Rebecca Haugen
RJH98	pIZ MS2 RD3 2/3 alone, 80-162	Rebecca Haugen
RJH101	pIZ MS2 RD3 ▲FA1+4	Rebecca Haugen
RJH115	pIZ MS2 RD3 1-195	Rebecca Haugen
RJH116	pIZ MS2 RD3 73-195	Rebecca Haugen
RJH135	pIZ MS2 RD3 CR4 NN184-185AA	Rebecca Haugen
RJH136	pIZ MS2 RD3 CR4 MF186-187AA	Rebecca Haugen
RJH137	pIZ MS2 RD3 CR4 GS188-189AA	Rebecca Haugen
RJH138	pIZ MS2 RD3 CR4 NS190-191AA	Rebecca Haugen
RJH139	pIZ MS2 RD3 CR4 SIF192-194AAA	Rebecca Haugen
RJH146	pIZ MS2 RD3 CR4 F187A	Rebecca Haugen
RJH147	pIZ MS2 RD3 CR4 F194A	Rebecca Haugen

Archive number	Plasmid name	Creator
RJH148	pIZ MS2 RD3 CR1 SLTG-A's	Rebecca Haugen
RJH149	pIZ MS2 RD3 CR1 RRD-A's	Rebecca Haugen
RJH150	pIZ MS2 RD3 CR1 SFD-A's	Rebecca Haugen
RJH151	pIZ MS2 RD3 CR1 RST-A's	Rebecca Haugen
RJH152	pIZ MS2 RD3 CR1 SAF-A's	Rebecca Haugen
RJH153	pIZ MS2 RD3 CR1 SPS-A's	Rebecca Haugen
RJH158	pIZ MS2 eGFP MCS	Rebecca Haugen
RJH161	pIZ MS2 RD3 CR4 F187+194A	Rebecca Haugen
RJH162	pIZ MS2 RD3 CR4 SSS-A's	Rebecca Haugen
RJH163	pIZ MS2 RD3 CR1 SLTG+CR4 F187A	Rebecca Haugen
RJH164	pIZ MS2 RD3 CR1 RRD+CR4 F187A	Rebecca Haugen
RJH165	pIZ MS2 RD3 CR1 SFD+CR4 F187A	Rebecca Haugen
RJH166	pIZ MS2 RD3 CR1 RST+CR4 F187A	Rebecca Haugen
RJH171	pIZ MS2 RD3 49-195	Rebecca Haugen
RJH172	pIZ MS2 RD3 CR1 SPS+CR4 F187A	Rebecca Haugen
RJH173	pIZ MS2 RD3 CR1 SAF+CR4 F187A	Rebecca Haugen
RJH240	pIZ MS2-eGFP-link-CR1-link-CR4 (1x)	Rebecca Haugen
RJH241	pIZ MS2-eGFP-link-CR1-link-CR4 (2x)	Rebecca Haugen
RJH242	pIZ MS2-eGFP-link-CR1-link-CR4 (3x)	Rebecca Haugen
RJH243	pIZ MS2-eGFP-link-CR1-link-CR4 (4x)	Rebecca Haugen
RJH184	pIZ MS2-eGFP-link-CR1-link-CR4 (5x)	Rebecca Haugen
RJH185	pIZ MS2-eGFP-link-CR1-link-CR4 (6x)	Rebecca Haugen
RJH186	pIZ MS2-eGFP-link-CR1-link-CR4 (5x) FF mut	Rebecca Haugen
RJH187	pIZ MS2-eGFP-RD3	Rebecca Haugen

Table 2.1: Plasmids used in Chapter 2. A full list of plasmids can be found in Appendix C.

Antibodies

All antibodies used in this work are listed in Appendix B. Antibodies used in this section are as follows: Mouse anti-V5 primary antibody (Invitrogen, Cat# R960-25) was diluted 1:5000 in blotto. Mouse anti-Tubulin (CST, Cat# 3873) was diluted 1:1000 in blotto. Rabbit anti-Vinculin (Thermo Fisher Cat# 700062) was diluted 1:1000 in blotto. Goat anti-rabbit-HRP secondary antibody (CST, Cat# 7074P2 or Sigma Cat# AP187P) was diluted 1:5000 in blotto. Goat anti-mouse-HRP (Thermo Fisher Cat# 31430) was diluted 1:5000 in blotto.

Chapter 3: Interactions of RD3 with the CNOT complex

3.1. Introduction

The CNOT complex has a long-standing role as a mediator of decay, and as such, is known to interact with several other decay factors. Through Not9, the CNOT complex interacts with GW182 proteins and is thus recruited to decay transcripts targeted by micro RNAs (Braun, Huntzinger et al. 2011, Fabian, Cieplak et al. 2011, Chen, Boland et al. 2014). Tis11, an RBP that binds to AU-rich elements in the 3'-UTR of targets, interacts directly with the Not1 N-terminus (Fabian, Frank et al. 2013). The CNOT complex also recruits the RNA helicase Me31b (DDX6 in humans), a necessary factor for unwinding RNA structures preceding decay (Mathys, Basquin et al. 2014, Ozgur, Basquin et al. 2015). Through Me31b, the CNOT complex is additionally linked to the decapping complex members Dcp1, Dcp2, and decapping enhancers Edc3 and Ge-1 (EDC4) (Fromm, Truffault et al. 2012, Sharif, Ozgur et al. 2013). In turn, Dcp1 directly couples decapping to 5' exonucleolytic decay through a direct interaction with the ribonuclease Xrn1 (Braun, Truffault et al. 2012). The CNOT complex can also interact directly with Xrn1 (Chang, Muthukumar et al. 2019).

Early research in the developing fly embryo showed that regulation of hunchback RNA by Pum and nanos occurred through poly(A) tail shortening, implicating a deadenylase complex (Wreden, Verrotti et al. 1997). Subsequent studies in yeast demonstrated that PUF proteins also relied on the presence of Ccr4 and Pop2 and that Puf5f could directly bind to Pop2 (Goldstrohm, Hook et al. 2006, Goldstrohm, Seay et al. 2007). *Drosophila* Pum RBD was shown to interact *in vitro* with Pop2, and the human homologs CNOT7 and CNOT8 (Kadyrova, Habara et al. 2007). The RBD also promotes deadenylation (Weidmann and Goldstrohm 2012). Additionally, CCR4 interacts with Pum *in vivo* (Joly, Chartier et al. 2013). Human Pum proteins depend on the CNOT complex for activity, and associate with CNOT 7/8 and CNOT 6/6L *in vivo* (Van Etten, Schagat et al. 2012). Finally, it was demonstrated recently that *Drosophila* Pum could interact *in vitro* with multiple members of the human CNOT complex. This interaction is likely strengthened by multiple points of contact, since each one of Pum's RDs was able to interact with the CNOT complex modules (Arvola, Chang et al. 2020).

The activity of Pum RD3 is also dependent on the CNOT complex, and depletion of either Not1 or Pop2 eliminates repression by RD3 (Arvola, Chang et al. 2020). Additionally, immunoprecipitation experiments with overexpressed Flag-tagged Pum N-terminus were able to pull down endogenous Not1. This interaction was not disrupted by the addition of RNase, indicating that it could be due to a direct protein contact. *In vitro* assays using MBP-tagged RD3 were able to pull down the human CNOT complex members, but it is unclear if the interaction is direct with a single complex member, or via multiple subunits (Arvola, Chang et al. 2020).

Pum activity has also been tied to decapping. Yeast Puf5p interacts with Eap1p, a 4E-BP, which interferes with translation and recruits the decapping protein Dhh1 (Blewett and Goldstrohm 2012). In *Drosophila* and human cells, the repressive activity of the Pum N-terminus is reduced when decapping is inhibited (Arvola, Chang et al. 2020, Enwerem, Elrod et al. 2021).

We hypothesized that RD3 was directly recruiting the decay machinery through conserved motifs. We set out to test the interactions between RD3 and individual CNOT complex members and interactions between RD3 and other decay factors.

3.2. Interaction screens - Y2H assays

To determine the CNOT complex members that RD3 may be associating with, we turned to the yeast 2-hybrid system (Y2H). RD3 was expressed as a Gal4 binding domain fusion (BD-RD3) on a plasmid carrying the gene for leucine synthesis (pGBK), and all potential interactors were cloned as activating domain fusions (AD-Nots) with the gene for tryptophan synthesis (pGAD). Both plasmids were co-transformed into Y2H gold strain yeast, auxotrophic for adenine, histidine, leucine, and tryptophan. Adenine and histidine synthesis occurred only when the tested proteins were able to functionally interact. Additionally, interactions activate transcription of the MEL1 gene, which encodes the protein alpha-galactosidase, and enables synthesis of the chromogenic substrate x-alpha-gal, turning the colonies blue. These three reporter genes are under the control of the GAL4 UAS promotor. Co-transformants in which the proteins do not interact also produce a red pigment due to accumulation of an intermediate in the adenine biosynthesis pathway.

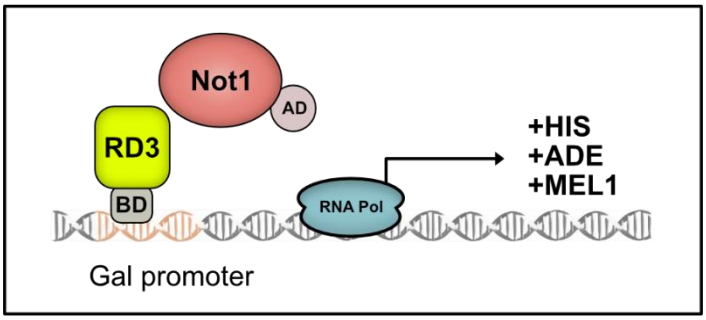


Figure 3.1: The yeast two hybrid system. RD3 constructs are fused to the DNA binding domain (BD) and Not1 or other proteins are fused to the activation domain (AD). The plasmids expressing these constructs allow yeast transformed with both to grow on media lacking tryptophan and leucine. When the two proteins interact, transcription of reporter genes begins, allowing growth on media also lacking adenine and histidine. Transcription of the MEL1 gene allows yeast to synthesize a chromogenic substance which turns the colonies blue.

Since knockdown of Not1 and Pop2 had the greatest effect on repression by RD3 (Arvola, Chang et al. 2020), we began by testing if Not1 could interact with RD3 *in vivo* in yeast. The large size of Not1 necessitated the creation of 3 separate regions: the N-terminal region (Not1-N, residues 1 to 1148), the MIF4G region (Not1-M, residues 1147 to 1717), and the C-terminal region (Not1-C, residues 1710 to 2503) (Figure 3.2). Each separate region of Not1 is capable of adopting stable structures *in vivo* and associating with its respective complex members (Petit, Wohlbold et al. 2012, Bawankar, Loh et al. 2013).

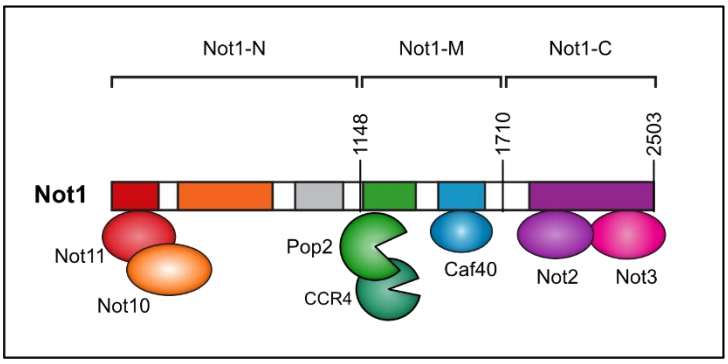


Figure 3.2: The arrangement of CNOT complex modules and Not1 subdivisions. Not1 serves as a scaffold for the other subunits to bind. The subdivisions of Drosophila Not1 used in this work are shown above the diagram with corresponding amino acid positions. Each subdivision of Not1 is capable of associating with the respective subunits in each module (Petit, Wohlbold et al. 2012). Figure adapted from (Chang, Muthukumar et al. 2019).

RD3 interacted with Not1-N, but not Not1-M or Not1-C (Figure 3.3). This was not due to auto-activation of either RD3 or Not1, as neither construct was capable of growth when transformed with empty vector.

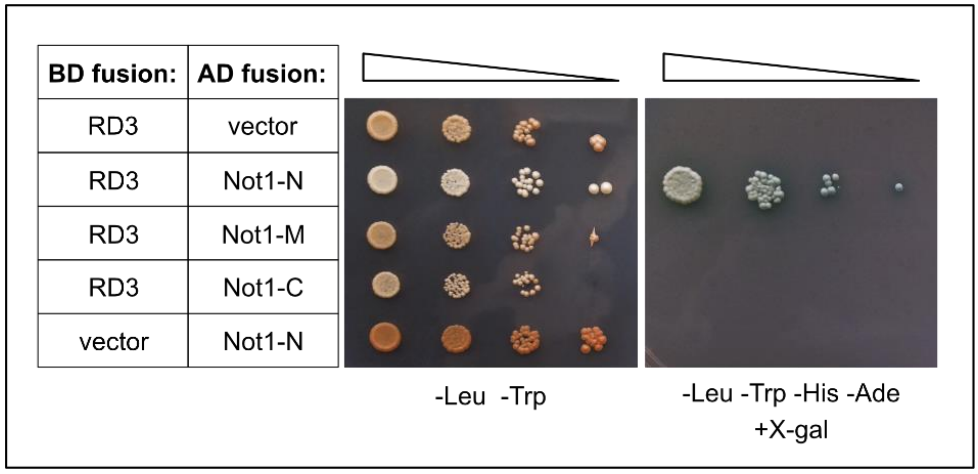


Figure 3.3: RD3 interacts with Not1-N in yeast. (Left) Co-transfected plasmids are listed. The binding domain (BD) is expressed with RD3 proteins or empty vector expressing only the BD. The activating domain (AD) is expressed with the Not1 constructs, or alone (vector). Vector controls ensure the fusion protein is not capable of either binding or activating without a protein interaction with the other domain fused to the tested protein (auto-activation). (Middle) Photograph of yeast serial dilutions (10^4 to 10^1) plated on media lacking leucine and tryptophan (DDO). DDO media selects for colonies containing both AD and BD plasmids. (Right) Photograph of yeast serial dilutions duplicate plated on media lacking leucine, tryptophan, histidine, and adenine, (QDO) and supplemented with the chromogenic substrate X-α-gal. Colony growth on this plate indicates a positive interaction between the tested proteins. All Y2H figures in this chapter are illustrated in this manner. n ≥ 3.

Other CNOT complex members that affected RD3 activity were Not2 and Not3 (Arvola, Chang et al. 2020). These two subunits, along with the C-terminal region of Not1, form the primary Not module. The Not module mediates complex assembly in cells and serves as an interaction platform for multiple regulatory proteins including Nanos (Maillet, Tu et al. 2000, Boland, Chen et al. 2013, Bhandari, Raisch et al. 2014). Since the Pum RBD binds to and interacts with Nanos (Weidmann, Qiu et al. 2016), it was reasoned that RD3 might also interact with the Not module as well. Since an interaction between RD3 and Not1-C was not seen, we tested Not2 and Not3. Full-length Not3 (aa 1 to 844) tested with RD3 was negative (data not shown), so we instead tested a

truncation of Not3 encompassing the C-terminal NOT box domain, aa 522 to 844 (Not3-C), and observed that RD3 can interact with Not2 and Not3-C (Figure 3.4).

Pop2 is also necessary for RD3's activity (Arvola, Chang et al. 2020). Surprisingly, there was no detectable interaction between RD3 and Pop2 (Figure 3.4). This implies that RD3's reliance on Pop2 for repressive activity is due solely to the enzymatic function of Pop2, since CCR4, the second deadenylase of the CNOT complex, is believed to be inactive in *Drosophila*. Additionally, RNAi depletion of CCR4 in Dmel2 cells did not affect the function of RD3 or any other RD of Pum (Arvola, Chang et al. 2020). CCR4 was also screened for a potential interaction with RD3 and the result was negative.

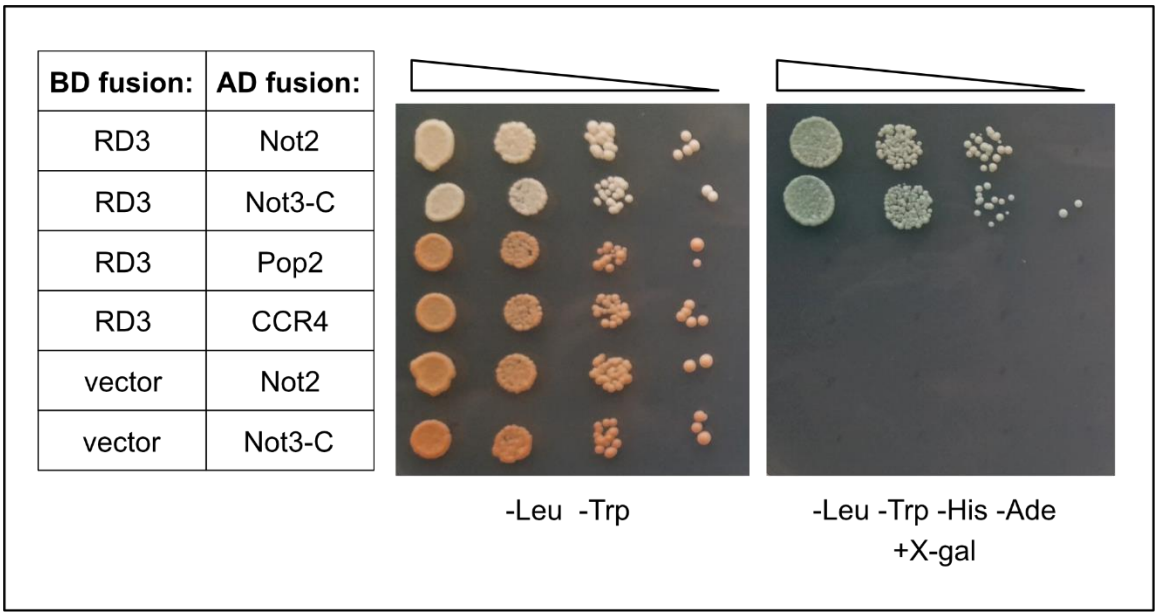


Figure 3.4: RD3 interacts with Not2 and Not3-C. BD and AD fusion proteins are listed, along with the negative controls, BD-vector with Not2 and Not3. As seen in the QDO plate on the right, RD3 interacts with Not2 and Not3-C, but not Pop2 or CCR4. n ≥ 3.

The other main CNOT complex members in fly were also tested for interactions with RD3: Caf40 (Not9), Not10, and Not11. Caf40 binds to the Not1-M region of Not1 (Bawankar, Loh et al. 2013). Neither Caf40 nor Not1-M appeared to interact with RD3. Not10 and Not11 both bind together at the Not1 N-terminus (Bawankar, Loh et al. 2013), which had shown a positive interaction with RD3, however neither protein appeared to stably interact with RD3 (Figure 3.5).

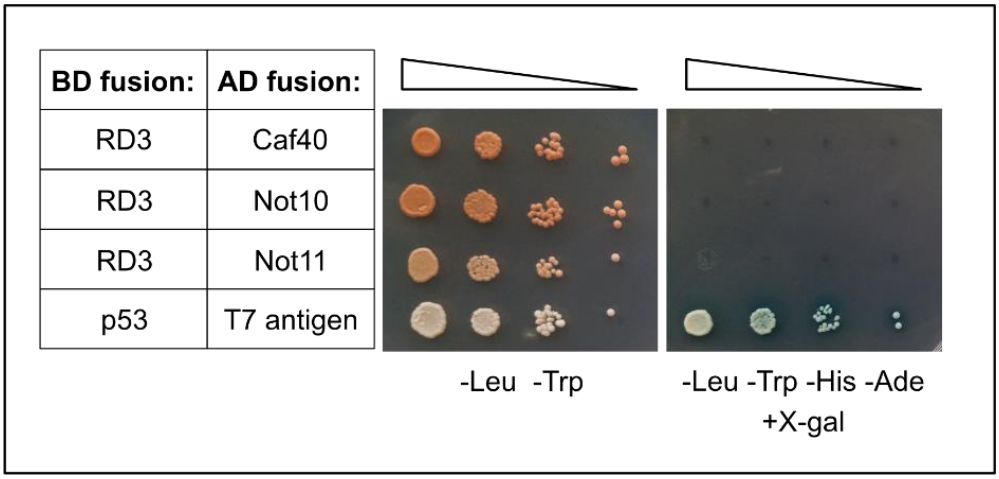


Figure 3.5: RD3 does not interact with Caf40, Not10, or Not11. BD and AD fusion proteins are listed, along with the assay positive control p53+T7 antigen. As seen in the QDO plate on the right, none of the tested CNOT subunits interact with RD3. n ≥ 3.

Our initial assays provided evidence for a positive interaction between Not11 and Pum RD. However, through experimental replication this observation proved to be inconsistent. When multiple independent colonies from freshly transformed yeast were tested, the Pum RD3 – Not11 interaction was not reproducible (Figure 3.6). Not11 expression in these tests remained consistent, ruling out variability in expression levels of Not11. We also considered the possibility that variability might arise if *Drosophila* Not11 could associate with the endogenous yeast CNOT complex. Two facts argue against this notion. First, yeast do not possess a Not11 homolog. Second, yeast Not1 lacks the N-terminal region responsible for Not11 binding (Basquin, Roudko et al. 2012, Bawankar, Loh et al. 2013). The preponderance of evidence indicates that the RD3 and Not11 interaction in our initial two-hybrid assay represented a false positive.

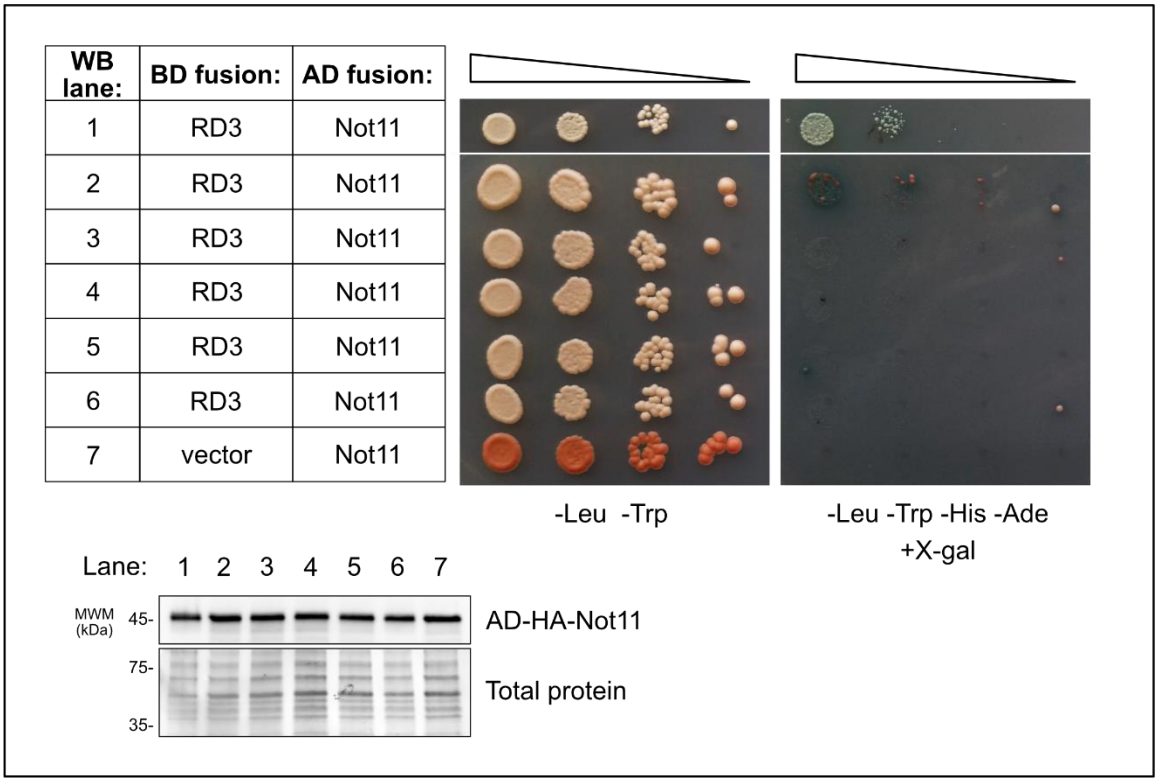


Figure 3.6: RD3 false positive interaction with Not11. (Top) A positive interaction was seen with Not11, but tests of multiple colonies (below) failed to grow. (Bottom) Western blot showing even expression of all Not11 constructs extracted from yeast. Total protein from Sypro Ruby stain is shown as a loading control.

Given the complex interplay of decay factors involving the CNOT complex, we decided to test if RD3 could interact with any potential accessory proteins as well. In addition, the positive interaction seen between RD3 and Not1-N also made us question whether RD3 could interact with Tis11. Tis11 is an ARE-binding protein and recent evidence shows enrichment of AREs surrounding functional PREs (Wolfe, Schagat et al. 2020). Me31b, Dcp1, and Tis11 were cloned into the pGAD vector and tested for interactions with RD3. All these results appeared negative, however the level of protein expression for Dcp1 and Tis11 were low. It is unclear if this level of protein would be enough to activate the transcription of genes if there were a positive interaction with RD3. Expression of all Y2H constructs is shown in Figure 3.14.

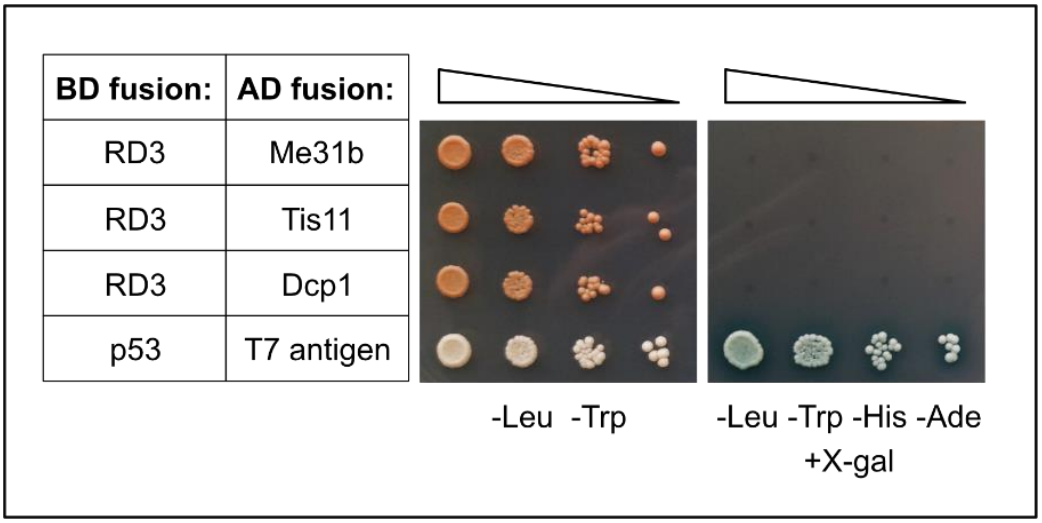


Figure 3.7: RD3 does not interact with Me31b, Tis11, or Dcp1. RD3 interactions tested with decay pathway accessory proteins fail to grow on QDO plates. The positive assay control p53+T7 antigen is shown at bottom. n ≥ 3.

Using the known structures of Not1 homologs, we decided to make truncations of Not1-N to determine where the interaction with RD3 was occurring (Figure 3.8). The first 412 residues of Not1 are involved in Not10/11 binding (Bawankar, Loh et al. 2013), and residues 908 to 1065 bind the RBP Tis11 (TTP in human) (Fabian, Frank et al. 2013). The N-terminus also contains multiple regions of HEAT repeats (Basquin, Roudko et al. 2012, Petit, Wohlbold et al. 2012, Fabian, Frank et al. 2013). We first divided the N-terminus into amino acids 1 to 788 and 789 to 1148. The first 788 residues did not interact with RD3, however, 789 to 1148 showed a positive interaction. Narrowing down the interaction further, we next tested amino acids 789 to 1064 and 1061 to 1148, then further subdivided the positive interaction between 789 to 1064 into 789 to 908, and the region of Tis11 binding, 908 to 1064. RD3 also interacted with the same region of Not1 that had been demonstrated to bind to Tis11 (Figure 3.8). This region for both human proteins TTP and CNOT1 is highly conserved in *Drosophila*. Interestingly, Fabian *et al.* also demonstrated that mutation of a single phenylalanine residue (F319) of TTP involved in binding to CNOT1 disrupted the interaction altogether *in vitro*. We further divided this 156 peptide-long region into 908 to 1047, 908 to 1016, and 908 to 995, however these smaller fragments were not expressed as well. Still, there was a positive interaction with all three fragments.

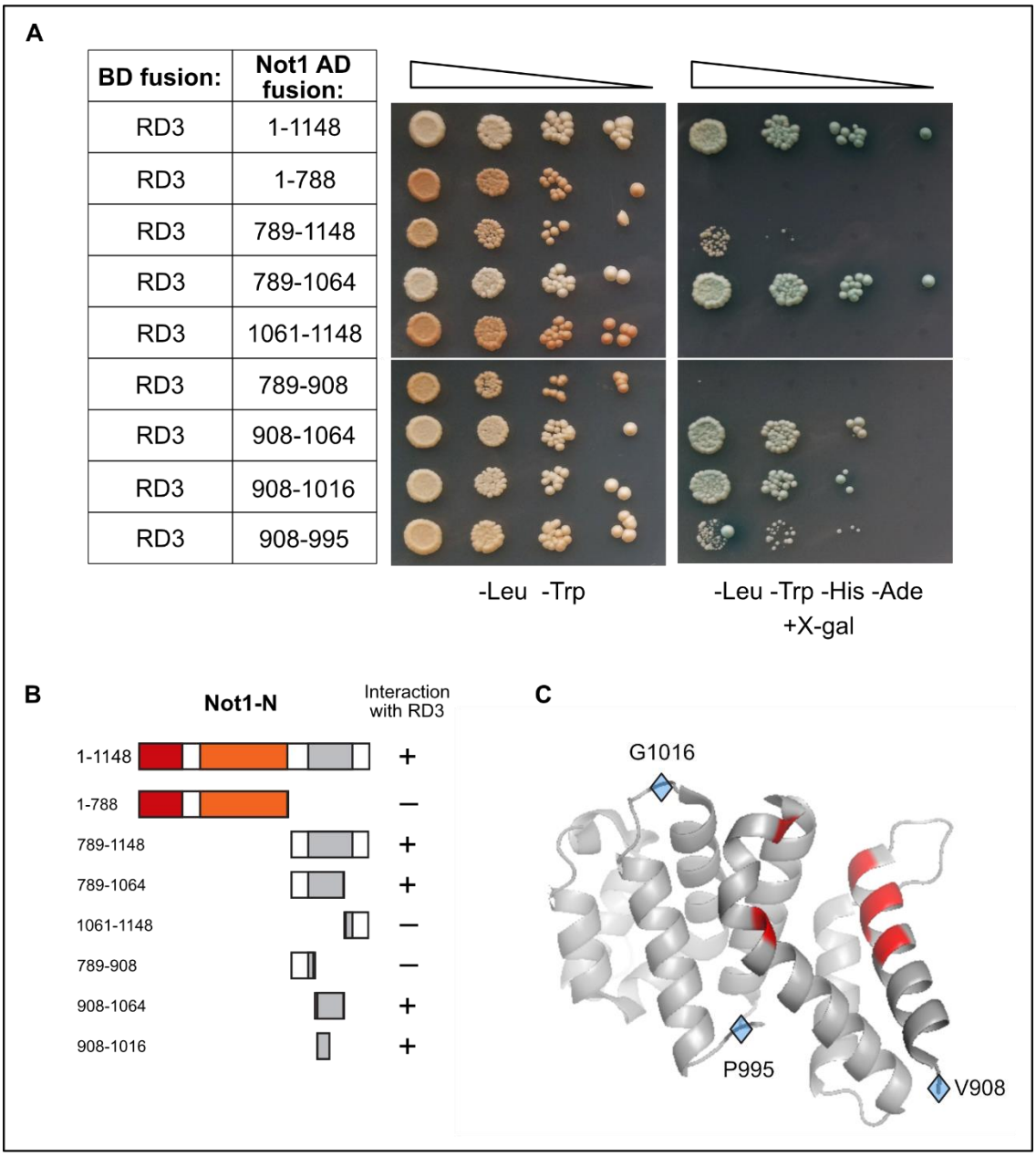


Figure 3.8: RD3 interactions with Not1 truncations. (A) Full-length RD3 was tested for interactions with truncations of Not1, shown by their amino acid numbers at left. The smallest regions of Not1 with a positive interaction are in the bottom panel. n ≥ 3. (B) Schematic representation of Not1 truncations tested for interaction with RD3. (C) Crystal structure of human Not1 aa 820 to 999 from (Fabian, Frank et al. 2013). Equivalent amino acid positions and numbering for *Drosophila* are shown as diamonds. The residues responsible for interaction with human TTP peptide are highlighted in red. Image generated with PyMOL, PDB 4J8S.

We employed the same strategy of protein truncations to determine which domain of Not2 RD3 was interacting with. Not2 is 585 amino acids in length and the structure for the C terminal region where it associates with the NOT module (Not1-Not2-Not3) is known for the human proteins (Boland, Chen et al. 2013, Raisch, Bhandari et al. 2016). This region is referred to as the NOT box domain, a feature that is shared between Not2 and Not3. In solution, both Not2 and Not3 adopt a similar fold and form a heterodimer that forms the binding interface with the C terminal region of Not1 (Figure 3.9). This crystallized region in human NOT2 corresponds to amino acids 399 to 585 in *Drosophila*.

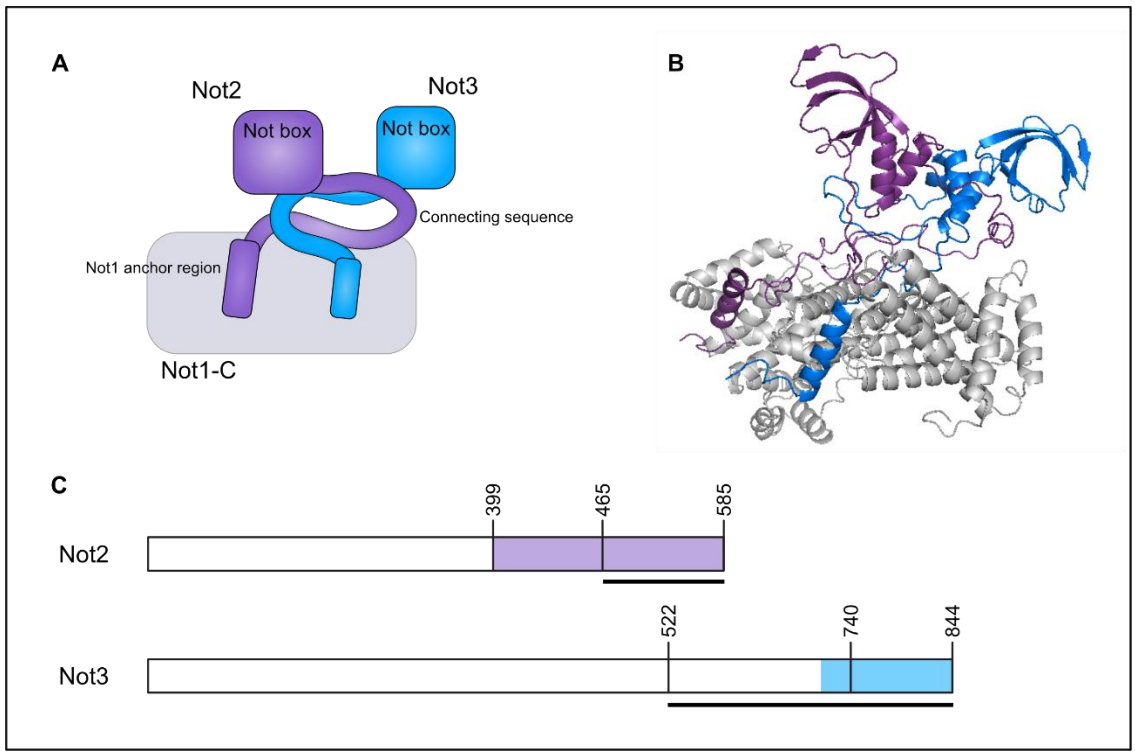


Figure 3.9: The components of the NOT module. (A) Cartoon representation of the NOT module structure shown in B. The C-terminal regions of Not1, Not2, and Not3 are shown. Not2 and Not3 have conserved C-terminal domains referred to as NOT boxes. A short peptide sequence called the connector sequence (CS) mediates the interaction between Not2 and Not3. The Not1 anchor region (NAR) on each protein binds to Not1. All three proteins have N-terminal regions that are not shown in A and B. (B) Crystal structure generated with PyMOL from PDB 5FU7 (Raisch, Bhandari et al. 2016). Proteins are colored as in A; Not1-C grey, Not2 purple, Not3 blue. Not2 and Not3 N-termini have not been crystallized. (C) Stick diagrams of Not2 and Not3 with amino acid positions for tested truncations shown above. Colored shading refers to known crystal structure from B. Truncations with positive RD3 interactions are underlined.

We tested Not2 constructs of aa 1 to 399 and 399 to 585 for interaction with RD3 in the Y2H system. A positive interaction was seen for the C terminal area of Not2 (399 to 585, Figure 3.10), therefore we further divided this area into aa 399 to 475 and aa 465 to 585. The area chosen for the division falls within the connector sequence, the region of interaction between Not2 and Not3 (Boland, Chen et al. 2013). This effectively divided Not2 into a section that binds to Not1 (Not1 anchor region, aa 399 to 475) and the remaining Not box domain (aa 465 to 585). RD3 interacted with Not2 aa 465 to 585 (Figure 3.10).

A similar truncation was made for Not3-C. The region of aa 522 to 844 was divided in the CS region to isolate the NOT box domain, aa 740 to 844. However, this peptide did not interact with RD3 in yeast (data not shown).

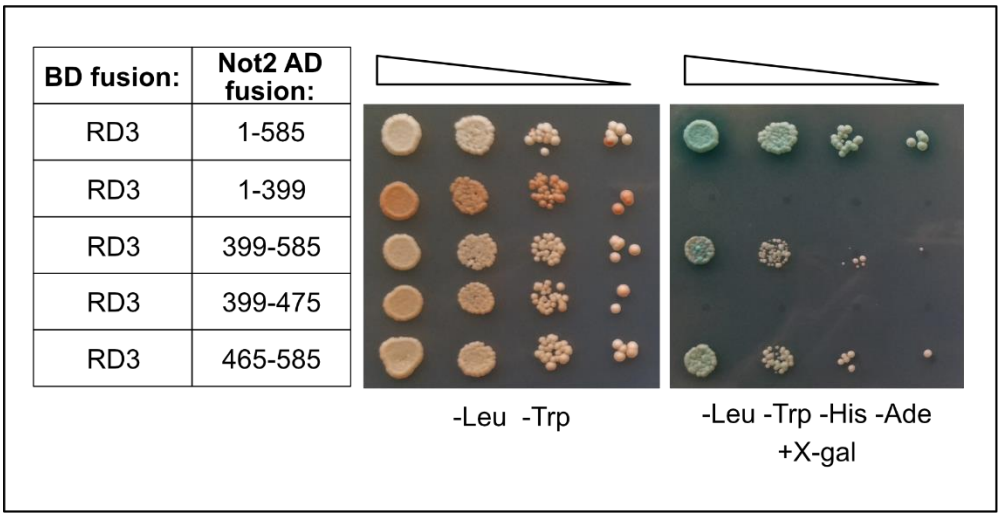


Figure 3.10: The interactions of Not2 subdivisions with RD3. The Not2 truncations are listed by their amino acid numbers at left. The C-terminal Not box domain shows a positive interaction with RD3 on QDO media. n ≥ 3.

Having identified the CNOT complex subunits that interact with RD3, we next tested if the phenylalanine mutations in CR4 would disrupt those associations in yeast. Mutating F1033 was enough to eliminate the interaction between RD3 and Not1-N (Figure 3.11). Additionally, the F1040A mutation also abolished the interaction between RD3 and Not1-N. Deleting CR1 affected the interaction with Not1-N as well, but this deletion alone was not enough to completely disrupt it.

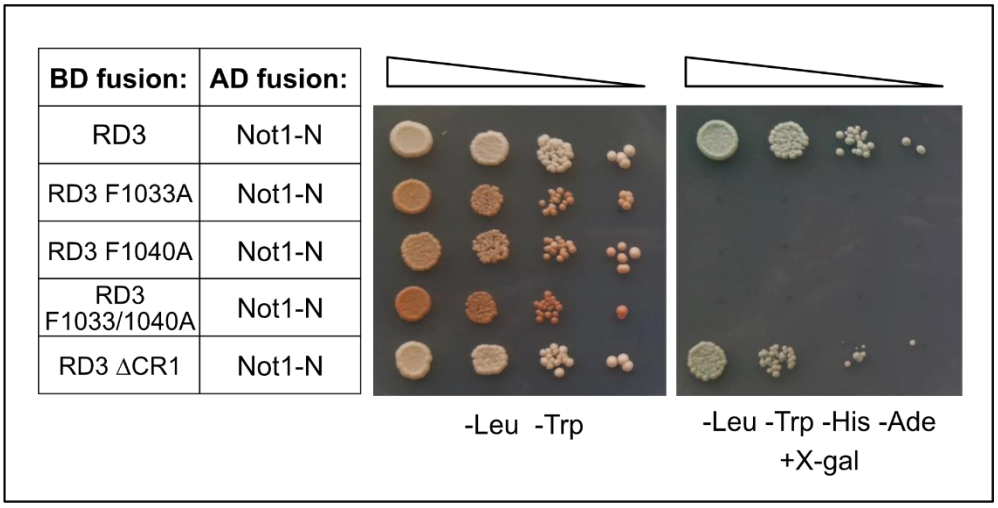


Figure 3.11: The effects of RD3 mutations on the interaction with Not1-N. The RD3 phenylalanine mutations and CR1 deletion are shown with Not1-N. Any single phenylalanine mutation disrupts the interaction with Not1-N, but not ΔCR1. n ≥ 3.

To test which areas of RD3 were responsible for the interaction with Not2 and Not3, we again tested the CR1 deletion and CR4 phenylalanine mutations. Deletion of CR1 had no effect on the Not2 interaction or Not3, nor did the single phenylalanine mutations (Figure 3.12). When both CR4 phenylalanines were mutated, the interactions with Not2 and Not3 were impacted. This data indicated that Not2 and Not3 were not contacting RD3 through CR1, but might be contacting CR4.

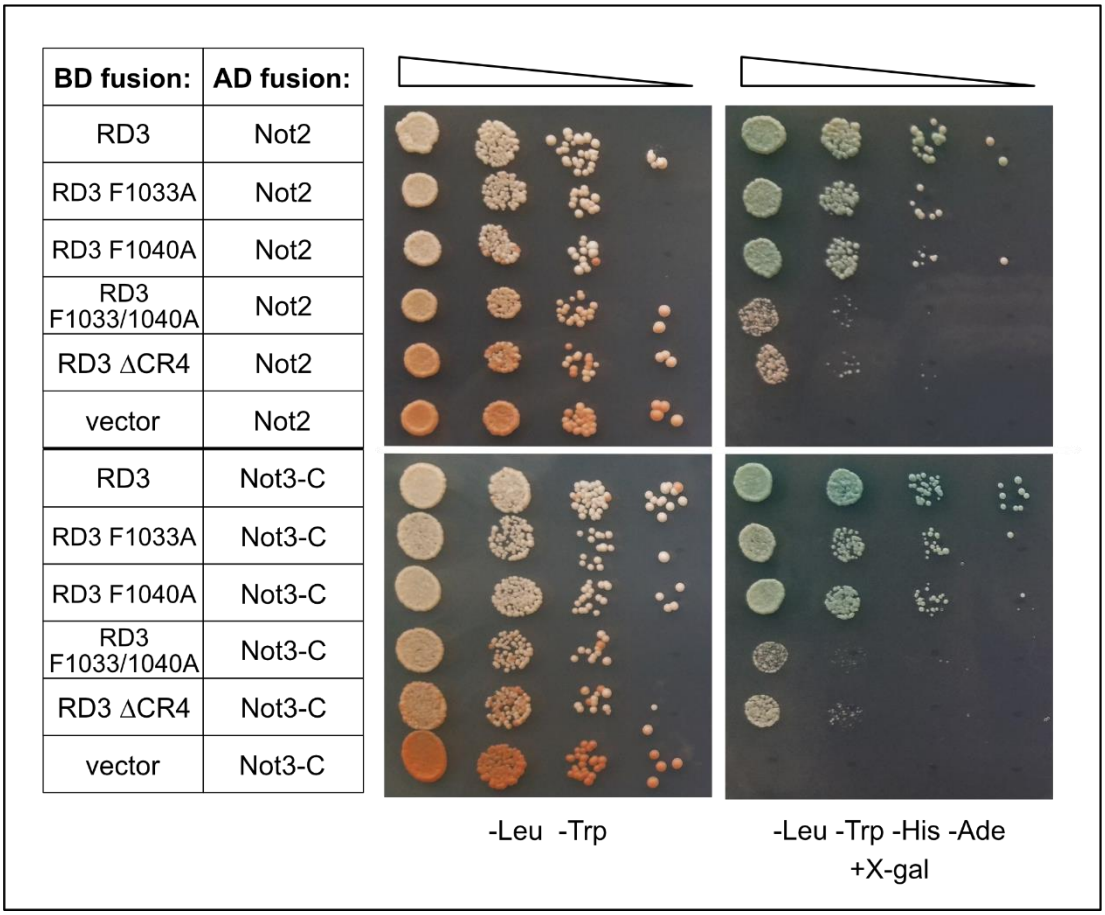


Figure 3.12: The effects of RD3 mutations on the interaction with Not2 and Not3-C. RD3 mutations were tested for interactions with Not2. Removing CR4 or mutating both phenylalanine residues affected the interaction but did not completely disrupt it. n ≥ 3.

Since deleting CR4 did not completely disrupt the interaction with Not2, next we screened deletions of the other CRs, since several of them showed a decrease in RD3's activity when deleted in our luciferase assay system. However, none of the other CRs appeared to be responsible for the interaction with Not2 (Figure 3.13). We tested the same RD3 deletions with Not3, and observed that deleting both CR1 and CR4 together disrupted the interaction between RD3 and Not3 (Figure 3.13).

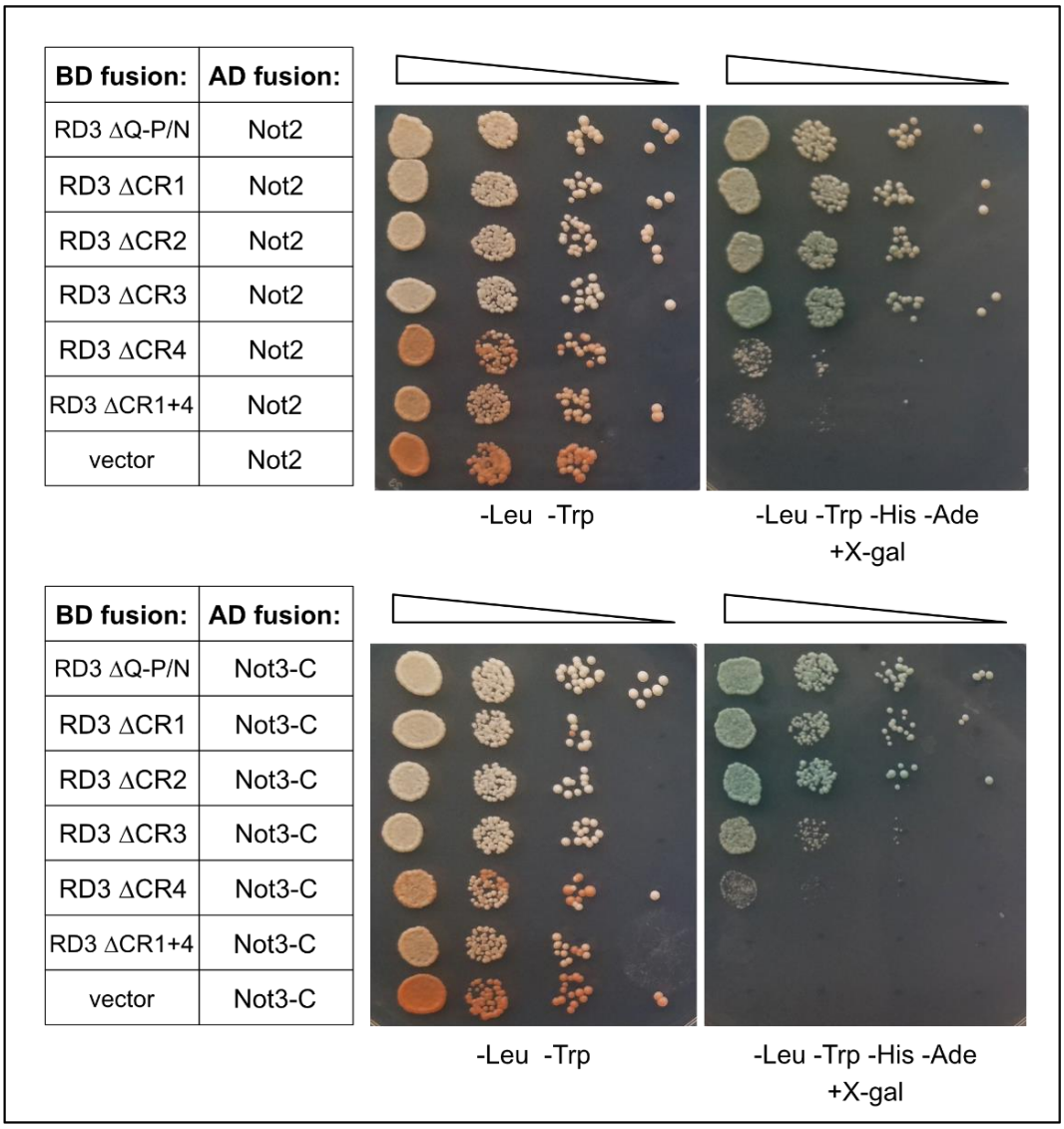


Figure 3.13: The effects of RD3 deletions on the interaction with Not2 and Not3. Deletions of all the CRs of RD3 were tested for interactions with Not2 (Top). None of the deletions completely disrupt the interaction. (Bottom) RD3 deletions tested with Not3. The double deletion of CR1 and CR4 disrupts the interaction. n ≥ 3.

We attempted to determine if the 5x CR1+4 constructs interact with Not1-N or Not2 using the Y2H system. However, the 5x CR1+4 construct auto-activated the reporter genes, so we were not able to test this interaction. Curiously, the 5x CR1+4 FF mutant did not show auto-activation, but did not interact with either Not1-N or Not2 (data not shown).

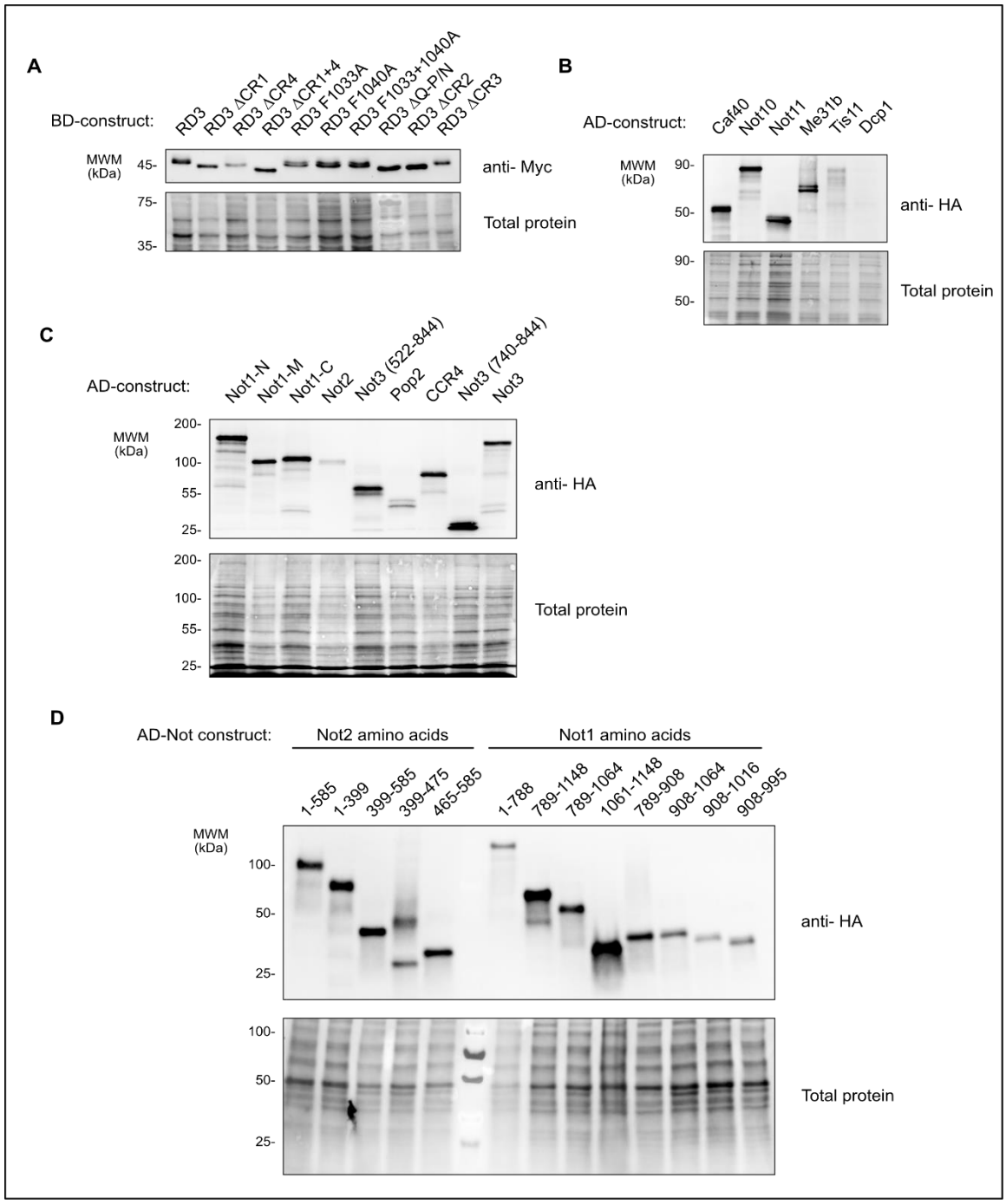


Figure 3.14: Expression data for Y2H constructs. Western blots showing expression of all tested activation domain (AD) and binding domain (BD) fusion proteins. Sypro Ruby stain was used as a loading control for total protein. (A) Myc- tagged BD-RD3 constructs. (B) HA-tagged AD constructs. Tis11 and Dcp1 were not expressed well. (C) HA-tagged AD-fusion CNOT complex members. The Not3 (522 to 844) construct is used in interaction tests with RD3. Not3 full-length was still expressed. (D) Truncations of HA-tagged Not2 and Not1-N AD fusions. Molecular weight marker (MWM) can be seen in the center of the gel.

3.3. Immunoprecipitation assays in *Drosophila* cells

To corroborate our findings from the Y2H system, we turned to immunoprecipitation assays (IP) in *Drosophila* cells. First, we tested the ability of RD3 and the CR4 double phenylalanine mutant, F1033A+F1040A (RD3 FF mut) to interact with Not1. We used a DL1 cell line expressing Flag-tagged Not1 which we generated with CRISPR-Cas9. For negative controls, we expressed RD3 constructs with wild-type DL1 cells (WT), and HaloTag (HT)-V5 with Not1-Flag. Our positive control was Not11-V5. Additionally, all samples in these experiments were treated with RNase One and RNase A. RD3 co-immunoprecipitated with endogenous Not1 whereas HT-V5 did not (Figure 3.15). As expected, Not11-V5 interacted with Not1. Intriguingly, the RD3 FF mut also associated with Not1-N.

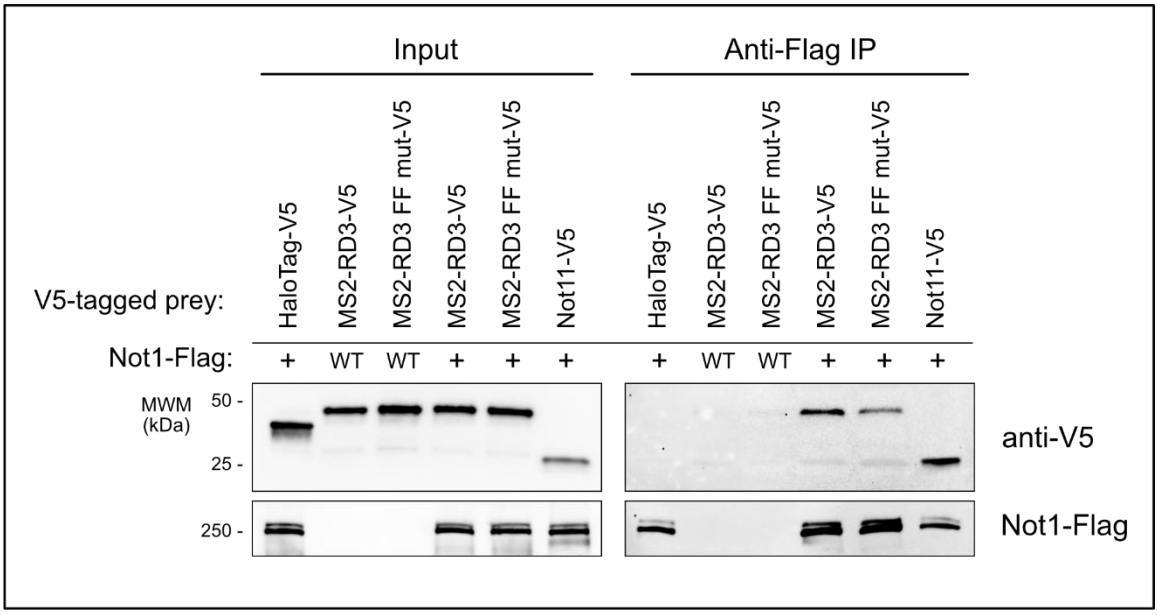


Figure 3.15: Immunoprecipitation assay with RD3 and endogenous Not1 in *Drosophila* cells. V5-tagged RD3 or RD3 F1033+F1040A mutant (RD3 FF mut) were transfected into Dl1 cells expressing endogenous Flag-tagged Not1, or WT DL1 cells. V5-HaloTag and Not11-V5 served as negative and positive controls, respectively. Expression of all samples is shown in the Input at left. (Right) Anti-Flag IP was used to capture interactions with RD3 and RD3 FF mut, and the positive control Not11. n ≥ 3.

Next, we co-expressed Flag-tagged Not1-N with wild type or mutant versions of V5-tagged MS2-RD3 in Dmel2 cells. HT-V5 served as a negative control and Not11-V5 was included as a positive control. The tested RD3 constructs were transfected with empty vector to rule out nonspecific interactions between RD3 and the Flag resin. RD3 co-immunoprecipitated with Flag-tagged Not1-N (Figure 3.16). The RD3 FF mut also co-immunoprecipitated with Not1, whereas the HT negative control did not.

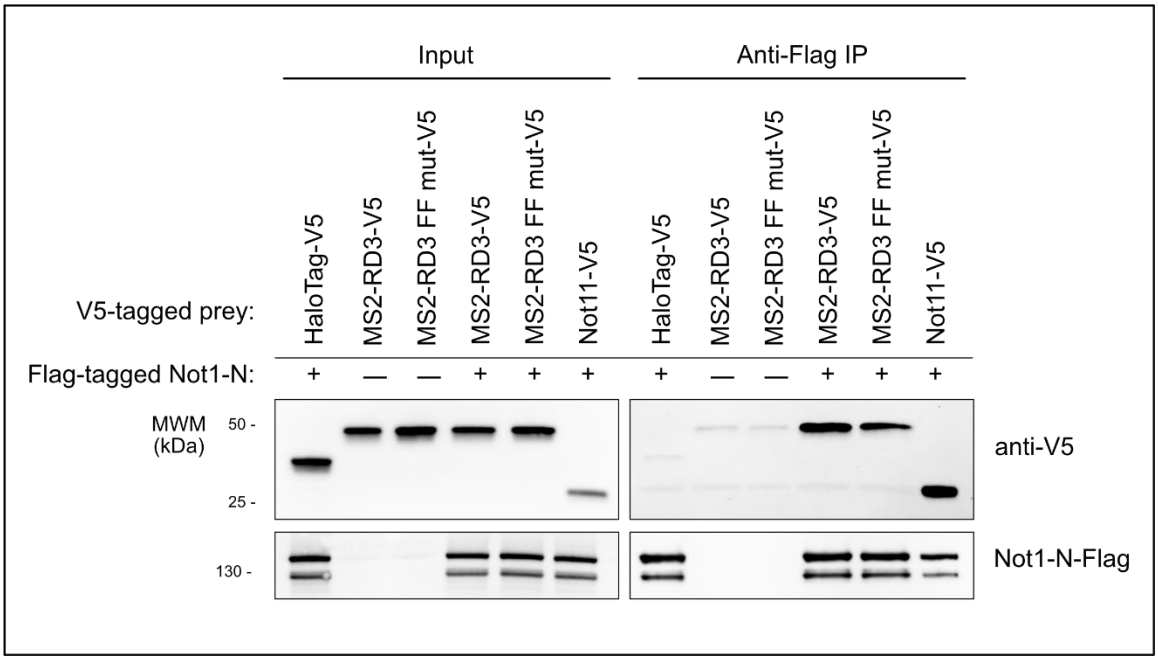


Figure 3.16: Immunoprecipitation assay with RD3 and Not1-N in *Drosophila* cells. V5-tagged RD3 or RD3 F1033+F1040A mutant (RD3 FF mut) were co-transfected with Flag-tagged Not1-N or an empty vector control. V5-HaloTag and Not11-V5 served as negative and positive controls, respectively. Expression of all samples is shown in the Input at left. (Right) Anti-Flag IP was used to capture interactions with RD3 and RD3 FF mut, and the positive control Not11. n ≥ 3.

Since our yeast data also indicated that RD3 can associate with Not2, we tested this interaction in Dmel2 cells as well. Using the same IP strategy of overexpressed V5 and Flag-tagged constructs, we were able to show that RD3 can interact with Not2. In this experiment Pop2 was used as a positive control with Not2 (Figure 3.17). The RD3 FF mutant also associated with Not2, in agreement with our observations in the Y2H system.

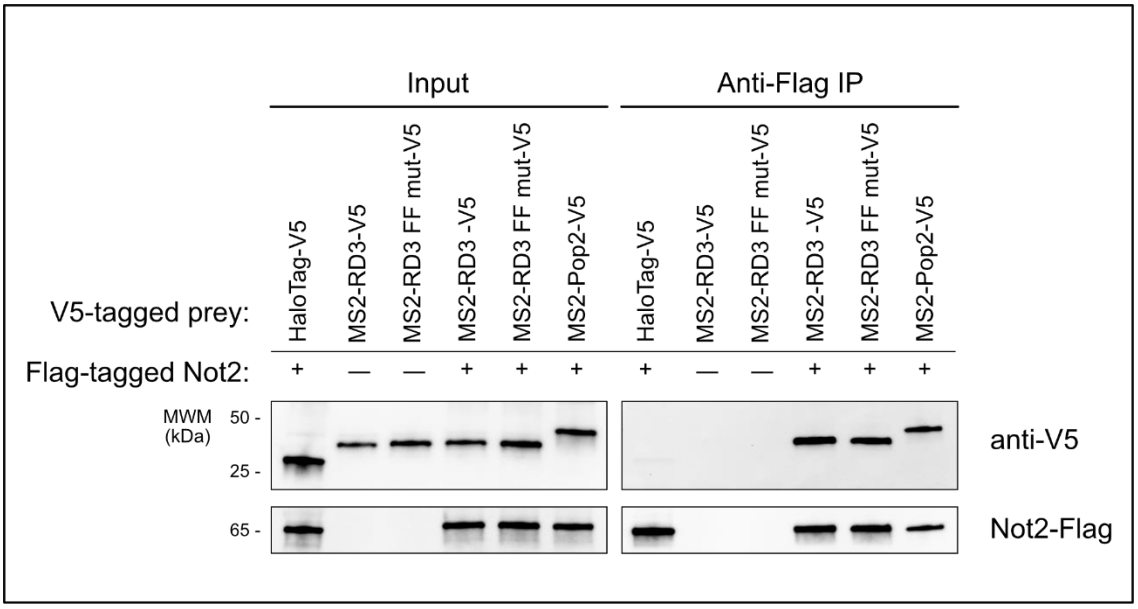


Figure 3.17: Immunoprecipitation assay with RD3 and Not2 in *Drosophila* cells. V5-tagged RD3 or RD3 F1033+F1040A mutant (RD3 FF mut) were co-transfected with Flag-tagged Not2 or empty vector control. V5-HaloTag and Pop2-V5 served as negative and positive controls, respectively. Expression of all samples is shown in the Input at left. (Right) Anti-Flag IP was used to capture interactions with RD3 and RD3 FF mut, and the positive control Pop2. n ≥ 3.

We next asked if the 5x CR1+4 construct could associate with the Not complex in *Drosophila* cells. Since the RD3 FF mutant could still interact with Not1 in cells, we expected that both the 5x CR1+4 and its mutant would be able to as well. We overexpressed the V5- tagged MS2-eGFP-5x CR1+4 construct (MG-5xCR1+4) and its mutant (MG-5xCR1+4 FF) with Flag-Not1-N in Dmel2 cells. We also included the same positive and negative controls as used previously. Both the 5x CR1+4 construct and its mutant interacted with Not1-N in cells (Figure 3.18 A).

One complication in this assay is the fact that the 5x CR1+4 FF mutant is more highly expressed than the WT version. We attempted to compensate for this discrepancy by transfecting less 5x CR1+4 FF mutant DNA, however, even at a 7.5:1 ratio of transfected plasmid (3 µg WT: 400 ng FF mut), the mutant version was still more abundant. Nonetheless, both constructs interacted with Not1-N.

We also tested whether the 5x CR1+4 multimer could interact with Not2 in Dmel2 cells (Figure 3.18 B). The 7.5:1 ratio of transfected DNA for the 5x CR1+4 constructs was used again. Both the 5x CR1+4 and the mutant co-immunoprecipitated with Not2.

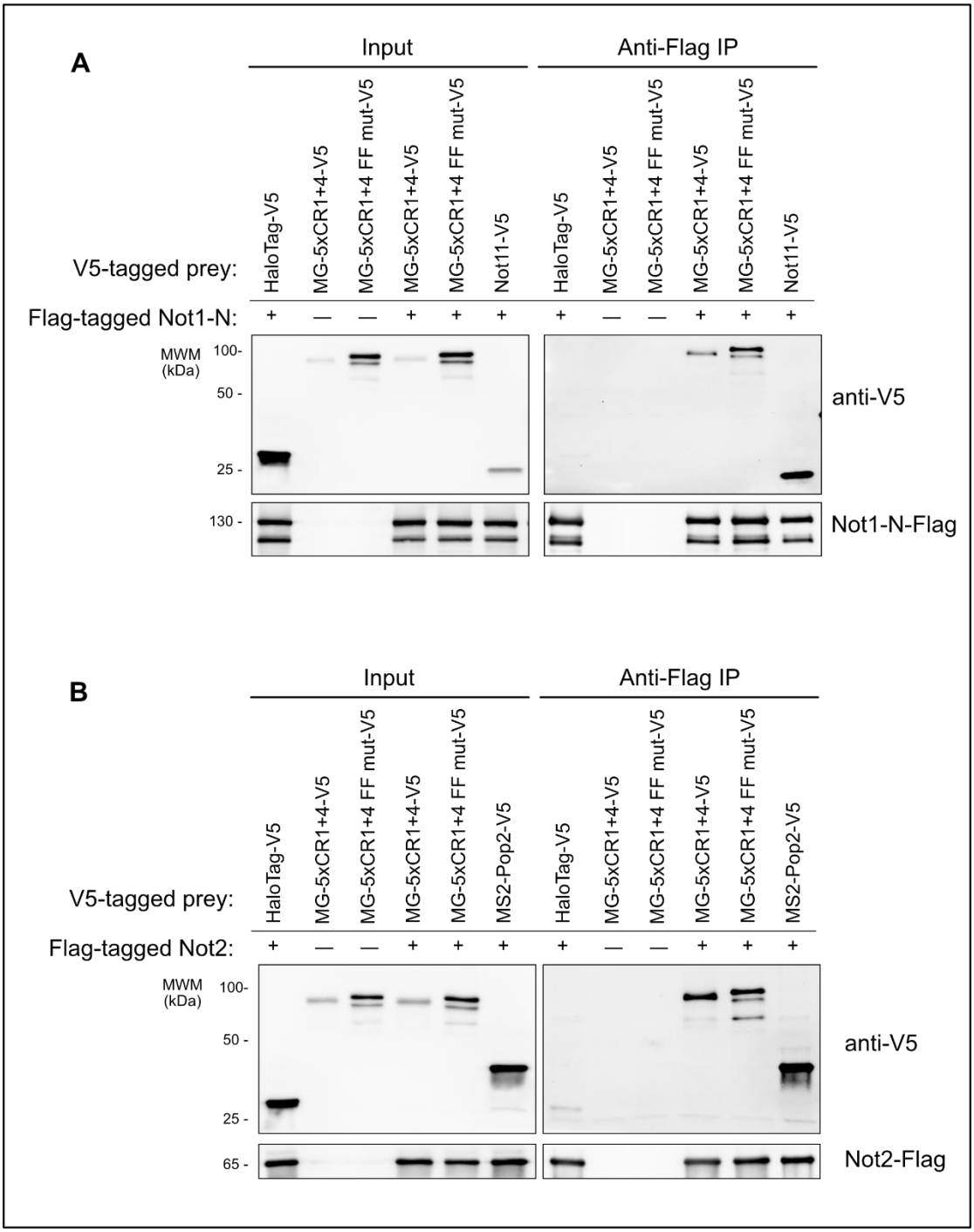
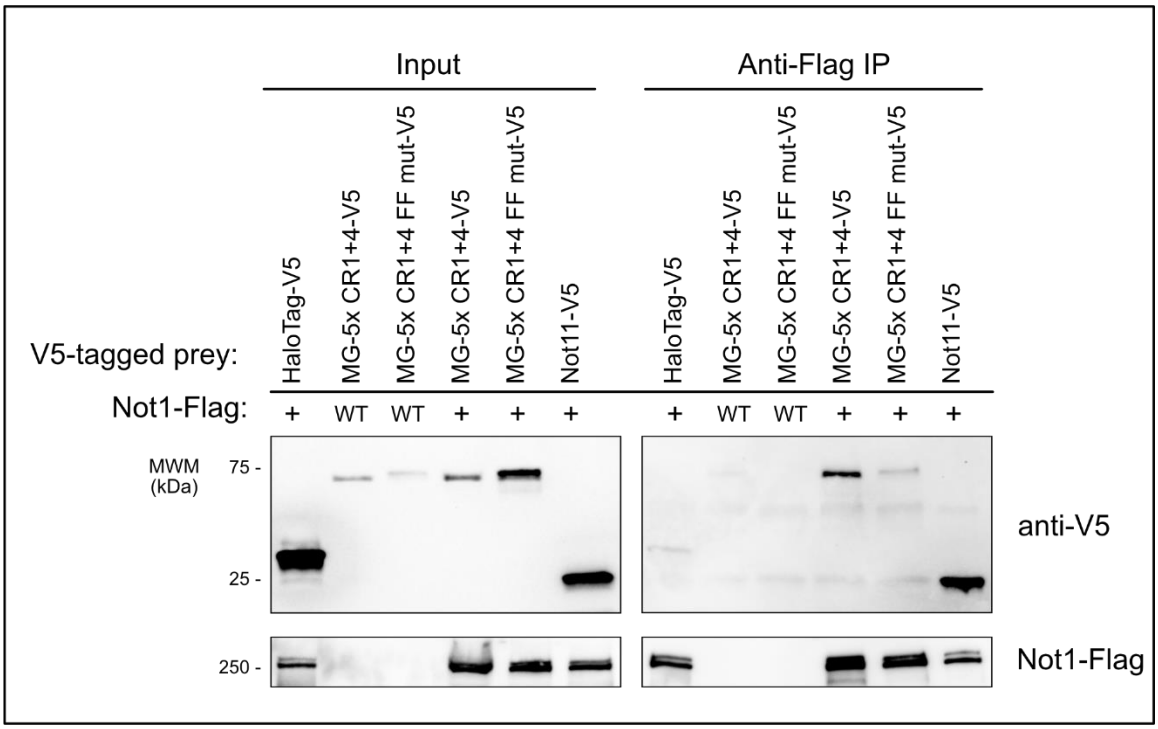


Figure 3.18: The 5x CR1+4 construct interacts with Not1-N and Not2 in *Drosophila* cells. The 5x CR1+4 construct and its mutant (5xCR1+4 FF mut) were overexpressed with Not1-N-Flag (A), Not2-Flag (B), or empty vector control (A, B). HaloTag-V5 served as the negative control as before. Not11 served as the positive control for Not1-N interactions, and Pop2 for Not2 interactions. n = 3.

Finally, we tested the interaction between 5x CR1+4 and its mutant in Dl1 cells expressing endogenous Not1-Flag. Because the protein expression levels of the 5x CR1+4 construct were extremely low in DL1 cells, we used three times the amount of total protein in both the WT and Not1-Flag IPs. The 5x CR1+4 FF mutants were still highly expressed in these cells, and we transfected 400 ng as in Dmel2 cells. Due to the inequality of total protein loaded into each IP we were unable to make comparisons between the 5x CR1+4 construct and its mutant, however, both the 5x CR1+4 and its mutant could interact with endogenous Not1 (Figure 3.19).



Fig**ure 3.19: The 5x CR1+4 and 5x CR1+4 FF mutant interact with Not1 in *Drosophila* cells.** The 5x CR1+4 construct and its mutant (5xCR1+4 FF mut) were overexpressed in either WT or Not1-Flag cells. HaloTag-V5 served as the negative control as before. Not11 served as the positive control for Not1 interactions. n = 3.

3.4. Summary and discussion

In this chapter, we set out to identify the protein interactions of Pum RD3 with the CNOT complex. We found that RD3 interacts with multiple members of the CNOT complex: Not1, Not2, and Not3. We also tested interactions between RD3 and other decay factors, Me31b, Tis11, and Dcp1, but did not detect any interactions. We mapped the regions of Not1 and Not2 that are necessary for these interactions. We found that RD3 interacted with Not1 in the N-terminal region, specifically, amino acids 908 to 1016. This is also the same region involved in TTP binding in humans (Fabian, Frank et al. 2013). RD3 interacted with Not2 in the conserved NOT box domain, aa 465 to 585. RD3 also interacted with Not3 in the C-terminal region, aa 520 to 844. We tested the effect of several mutations and deletions in RD3 that caused loss of function. For Not1-N, mutation of either phenylalanine residue in CR4 of RD3 destroyed the interaction. The interaction between RD3 and Not2 was resistant to the deletions of RD3's CRs, however the interaction between Not3 and RD3 was effected by deletion of CR1 with CR4.

Using co-IP, we demonstrated that RD3 interacted with the CNOT complex in *Drosophila* cells. RD3 co-immunoprecipitated endogenous Not1, and overexpressed Not1-N and Not2. This interaction was not disrupted by the CR4 double phenylalanine mutation for any of the CNOT members tested. We further demonstrated that the 5x CR1+4 multimer construct is sufficient to interact with the CNOT complex in *Drosophila* cells, and that mutating the phenylalanine residues in this construct also did not eliminate the interaction.

This was a surprising result since the phenylalanine mutations in CR4 destroyed RD3's repressive ability. A robust interaction between RD3 and the CNOT complex is assured through multiple contacts with Not1, Not2, and Not3. This interaction is evolutionarily conserved: in pull-down experiments with human proteins, RD3 interacts directly with all three subunits of the NOT module (Enwerem, Elrod et al. 2021). Since the physical interaction is not perturbed by the phenylalanine mutations, another mechanism must be responsible for the loss of RD3's activity.

RD3 might possess the ability to stimulate the CNOT complex's activity in addition to recruiting it to a transcript. The contacts with Not2 and Not3 might stabilize the interaction, while the functional phenylalanine residues might be responsible for a

stimulatory effect. There is precedence for fine-tuning of the CNOT complex's activity through RBPs. Other proteins that are known to affect the activity of the CNOT complex include Roquin (Sgromo, Raisch et al. 2017), Bag of Marbles (BAM) (Sgromo, Raisch et al. 2018), and Not4 (Raisch, Chang et al. 2019).

In vitro, a Roquin peptide inhibits deadenylation of a poly(A) substrate by competing with the substrate for binding to Caf40 (Raisch, Chang et al. 2019). In the proposed model, *in vivo* Roquin may block Caf40 binding to RNA and thus direct the CNOT complex deadenylase activity to the Roquin-bound transcript, therefore promoting specific target deadenylation versus non-specific "bulk" decay.

A similar mechanism may occur with RD3, where RD3 binds CNOT through Not2 and Not3 and directs the complex to decay the Pum-bound target RNA. While the RD3 FF mutant still associates with the complex via multiple contact sites with Not2 and Not3, without the stimulatory effect provided by the phenylalanines associating with Not1-N, decay is limited. An additional facet of this interaction between RD3 and Not1-N could be the exclusion of other RBPs such as Tis11 (TTP) which binds to Not1-N in the same region as RD3. This would provide a way for RD3 (and Pum) to compete with other RBPs for specific and rapid degradation of their own targets.

The only known structures of the entire complex come from yeast (Nasertorabi, Batisse et al. 2011, Ukleja, Cuellar et al. 2016). These electron microscopy structures depict the complex as L-shaped and lack the resolution needed to determine how RD3 might span the molecular distance between Not1-N and Not2. This structural prediction depicts the TTP binding face and the NOT module approximately 100 Å apart (Ukleja, Cuellar et al. 2016) and yet RD3 is predicted to be 43 Å in length. This implies that either RD3 does not contact both proteins simultaneously or that in cells, the CNOT complex adopts a very flexible structure. Indeed, it has been suggested that the CNOT complex does not form one single conformation, but rather multiple conformations, with each module acting like beads on a string (Raisch, Chang et al. 2019).

3.5. Materials and methods

Yeast two-hybrid assays

Yeast strain Y2H Gold from the Clonetech Matchmaker Gold kit (Takara Bio) were grown on YPAD (yeast extract peptone dextrose + adenine) agar plates. (Yeast genotype MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2 : : $GAL1_{UAS}$–$Gal1_{TATA}$–His3, $GAL2_{UAS}$–$Gal2_{TATA}$–Ade2 URA3 : : $MEL1_{UAS}$–$Mel1_{TATA}$).

Transformations of yeast were made using 1 µg total of both pGBK vector and pGAD vectors. Co-transformed yeast were grown on double dropout media (-Leu, -Trp, DDO) for 3 days at 30°C to select for co-transformants. For plate spotting, colonies from each DDO plate were struck into 5 mL DDO SD media and grown overnight at 30°C. The following day, 1 mL of culture was taken and grown in 4 mL YPAD media until the OD_{600} was approximately 1.0. Each culture was serially diluted from 2 x 10^6 cells/mL to 2 x 10^3 cells/mL and 5 µL of each dilution was duplicate plated on DDO and quadruple dropout (-Leu, -Trp, -His, -Ade, QDO) plates supplemented with 100 µL of 4 µg/mL alpha-x-gal substrate. Plates were incubated at 30°C for 3 to 5 days, then photographed. Each interaction was tested in at least three experimental replicates.

Western blotting for Y2H expression

For western blot detection, each co-transformant was grown overnight in 5 mL YPAD media as described. When cell density became ~OD_{600} 0.7, 10 million cells were harvested. Samples were prepared using 75 µL RIPA buffer (25 mM Tris pH 7.6, 1 mM EDTA, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) with 4x Roche cOmplete protease inhibitors (Sigma) and bead-bashed with glass beads on ice five times at 1 minute intervals. Additional 100x PMSF was added every 15 minutes. Sample lysate was cleared by centrifugation, 20 µL of 5x SDS loading dye was added, samples were incubated at 85°C for ten minutes, then 10 µL was loaded on a 12% SDS-PAGE gel, or 5-20% SDS-PAGE gradient gel (BioRad). Gels were electrophoresed at 120 to 150 volts until the dye front reached the bottom, then transferred to PVDF membranes (Immobilon). Total protein was detected by Sypro Ruby stain (Thermo Fisher) following the manufacturer's directions and imaged on a Chemidoc touch (BioRad). Primary antibodies were applied for one hour at room temperature or incubated overnight at 4°C. Membranes were washed three times for 10 minutes in blotto (1x PBS, 0.1% Tween, 5%

dried milk), then secondary antibodies were applied for 1 hour at room temperature. Membranes were washed three times for ten minutes each in blotto, then rinsed with 1x PBS before applying Pierce chemiluminescent substrate (Thermo Fisher). Blots were imaged using the auto exposure setting on a ChemiDoc touch (BioRad).

Co-immunoprecipitation assays

Dmel2 cells were seeded in 6 well plates at a density of 1.5 x 10^6 cells/mL in 2 mL SF900III media (Gibco). Transfection mixes were set up using 150 µL media, a total of 3 µg plasmid DNA (1.5 µg each for co-transfection; 3 µg for single transfection), and 12 µL FuGene HD (Promega). Transfection mixes for 5x CR1+4 differed in DNA amount as follows: for 5x CR1+4 FF mut, 400 ng 5x CR1+4 FF mut was transfected with empty pIZ vector balance to equal 1.5 µg total for co-transfection, or 3 µg total for single transfection in Not1-Flag cells. For 5x CR1+4, 4 µg total DNA was transfected with 16 µL FuGene HD for single transfection in Not1-Flag cells. Co-transfections in Dmel2 cells used 3 µg 5x CR1+4 and 1 µg Not1-N or Not2. Transfection mixes were incubated for 10 minutes before applying to cells. Cells were incubated at 25°C for 3 days before harvesting for co-IP assays.

Cells were collected and washed once with PBS, then incubated on ice for 10 minutes in lysis buffer containing 50 mM Tris pH 8.0, 500 nM NaCl, 1 mM EDTA, 0.2% Triton, and 2x protease inhibitors (Roche). Samples were mechanically disrupted with pestles for 20 seconds each, then lysates were cleared by centrifugation at 21,000 x g for 10 minutes. Lysate was passed through biospin columns (Millipore) and 35 µL of lysate was reserved for input analysis. Samples were then applied to prepared Flag resin.

20 µL of Flag resin suspension (EZview Red ANTI-FLAG M2 Affinity Gel, Sigma) was used per sample. Resin was prepared by washing four times in wash buffer containing 50 mM Tris pH 8.0, 500 nM NaCl, 1 mM EDTA, and collected by centrifugation at 4,000 x g between washes. Prior to binding, resin was resuspended in 700 µL per sample of lysis buffer with 1 µL per sample of both RNaseONE (Promega) and RNaseA (Promega.) Binding was performed with end-over-end rotation at 4°C for 2 to 3 hours.

After binding, the flowthrough was reserved and the samples were washed six times for 5 minutes each with end-over-end rotation at 4°C. The first three washes were performed with lysis buffer without protease inhibitors and the last three with wash buffer. After the final wash, approximately 20 µL volume remained of buffer and resin per sample.

Western blot imaging of IP assays

ELUATE: 20 µL of 2x SDS loading dye was added to each eluate sample and samples were incubated at 85°C for 10 minutes prior to loading on an SDS-PAGE gel. For dual over-expression samples, 10 µL of eluate was loaded for each sample, with the exception of positive CNOT controls, in which case 4 µL of eluate was loaded. For endogenous Not1 samples, 15 µL of eluate was loaded for all samples except positive controls, for which 5µL was used.

INPUT: Protein concentration of input samples was determined by Lowry DC protein assay (BioRad) and 5 µg of total protein (10 µg for endogenous Not1 samples) was loaded on either 12% or 5-20% gradient SDS-PAGE gel (BioRad). Gels were electrophoresed for approximately 1 hour at 150 to 180 volts, then transferred to PVDF membranes (Immobilon) at 4°C for 1.5 hours at 65 volts or overnight at 35 volts. Western blotting was performed by allowing the membrane to dry, then applying primary rabbit anti-V5 antibody for 1 hour with rocking at room temperature. Membranes were washed 3 times briefly in blotto (1x PBS, 0.1% Tween, 5% dried milk) then secondary goat anti-rabbit -HRP conjugated antibodies were applied for one hour. Membranes were washed twice for ten minutes each, then washed once with PBS before imaging. Chemiluminescent substrate was applied briefly (Pierce ECL, Thermo-Fisher or Immobilon, Sigma) and blots were imaged on the auto optimal exposure setting of a Chemidoc Touch (BioRad.) To image bait proteins, membranes were washed 3 times for ten minutes each in PBS, then primary rabbit anti-Flag antibody was applied to eluate blots, or mouse anti-Flag to input blots. Western blotting proceeded as described above.

Plasmids

Archive number	Plasmid name	Creator
RJH188	pGBK RD3	Rebecca Haugen
RJH189	pGBK RD3 ΔQ	Rebecca Haugen
RJH190	pGBK RD3 ΔCR1	Rebecca Haugen
RJH191	pGBK RD3 ΔCR2	Rebecca Haugen
RJH192	pGBK RD3 ΔCR3	Rebecca Haugen
RJH193	pGBK RD3 ΔCR4	Rebecca Haugen
RJH194	pGBK RD3 ΔCR1+4	Rebecca Haugen
RJH195	pGBK RD3 F1033A	Rebecca Haugen
RJH196	pGBK RD3 F1040A	Rebecca Haugen
RJH197	pGBK RD3 FF	Rebecca Haugen
CAW038	pIZ MS2CP-DmPop2 V5H6	Chase Weidmann
ACG544	pIZ GFP	Marzluf Lab
ACG272	pGADT7	CloneTech
ACG1036	pGADT7 T	CloneTech
ACG1037	pGBKT7 p53	CloneTech
RJH198	pGAD Not1-N	Rebecca Haugen
RJH199	pGAD Not1-C	Rebecca Haugen
RJH200	pGAD Not1-M	Rebecca Haugen
RJH201	pGAD Dcp1	Rebecca Haugen
RJH202	pGAD Me31b	Rebecca Haugen
RJH203	pGAD Tis11	Rebecca Haugen
RJH204	pGAD CCR4	Rebecca Haugen
RJH205	pGAD Pop2	Rebecca Haugen
RJH206	pGAD Not10	Rebecca Haugen
RJH207	pGAD Not11	Rebecca Haugen
RJH208	pGAD Not2	Rebecca Haugen
RJH209	pGAD Caf40	Rebecca Haugen
RJH210	pGAD Not3	Rebecca Haugen
RJH211	pGAD Not1-N 1-788	Rebecca Haugen
RJH212	pGAD Not1-N 789-1148	Rebecca Haugen
RJH213	pGAD Not1-N 789-1064	Rebecca Haugen
RJH214	pGAD Not1-N 1061-1148	Rebecca Haugen
RJH215	pGAD Not1-N 789-908	Rebecca Haugen
RJH216	pGAD Not1-N 908-1064	Rebecca Haugen
RJH217	pGAD Not1-N 908-1016	Rebecca Haugen
RJH218	pGAD Not1-N 908-995	Rebecca Haugen
RJH219	pGAD Not2 1-399	Rebecca Haugen
RJH220	pGAD Not2 399-585	Rebecca Haugen

Archive number	Plasmid name	Creator
RJH221	pGAD Not2 399-475	Rebecca Haugen
RJH222	pGAD Not2 465-585	Rebecca Haugen
RJH223	pGAD Not3 522-844	Rebecca Haugen
RJH224	pGAD Not3 740-844	Rebecca Haugen
pBA041	pGADT7 RD3	Brian Alzua
CAW018	pIZ HT-V5H6	Chase Weidmann
RJH175	pIZ Not1-N-Flag	Rebecca Haugen
RJH227	pIZ Not2-Flag	Rebecca Haugen
RJH228	pIZ Not11-V5	Rebecca Haugen
RJH184	pIZ MS2-eGFP-link-CR1-link-CR4 (5x)	Rebecca Haugen
RJH186	pIZ MS2-eGFP-link-CR1-link-CR4 (5x) FF mut	Rebecca Haugen
	pIZ MS2-RD3 V5 H6	Chase Weidmann
RJH161	pIZ MS2 RD3 CR4 F187+194A	Rebecca Haugen

Table 3.1: Plasmids used in Chapter 3. A full list of plasmids can be found in Appendix C.

Antibodies

All antibodies used in this study are also included Appendix B. Mouse anti-V5 primary antibody (Invitrogen, Cat# R960-25) was diluted 1:5000 in blotto. Rabbit anti-V5 (CST Cat# 13202S) was diluted 1:5000 in TBST. Mouse anti-Tubulin (CST, Cat# 3873) was diluted 1:1000 in blotto. Rabbit anti-Myc primary antibody (CST Cat# 2278S) was diluted 1:5000 in TBST. Rabbit anti-HA tag (CST Cat# 3724s) was diluted 1:5000 in TBST. Rabbit anti-Flag antibody (Sigma Cat# sab4301135) was diluted 1:5000 in blotto. Mouse anti-Flag (Sigma Cat# F3165-1MG) was diluted 1:5000 in blotto. Goat anti-rabbit-HRP secondary antibody (CST, Cat# 7074P2 or Sigma Cat# AP187P) was diluted 1:5000 in blotto. Goat anti-mouse-HRP (Thermo Fisher Cat# 31430) was diluted 1:5000 in blotto.

3.5.1. CRISPR generation of Not1-Flag Dl1 cells

Design of CRISPR guides

For all CRISPR workflows, the NCBI database was used to determine the ideal placement for CRISPR-tagging or gene knockout. If available, the N-terminus of each protein was chosen so that indels were more likely to result in a knockout. Benchling software (2019) was used to identify CRISPR sites with the highest potential efficiency

for Cas9 cleavage with the lowest potential off-target effects. Gene-specific primers (listed in Appendix G) were used to amplify the region of interest and TA-cloned into the pGEM T-easy vector to generate a library. Ten clones from each library were mini-prepped using Promega Mini-prep kits and sequenced to verify the targeted region matched the NCBI sequence and that the PAM sites were correct. Single-stranded DNA oligos were ordered to match the sgRNA sequence with additional BspQI overhangs for cloning into the pAc-sgRNA-Cas9 vector (Bassett, Tibbit et al. 2014). Oligos were prepared for cloning by taking 50 pmol of each forward and reverse strand, briefly heating them at 90°C, and then slowly cooling them. The annealed oligos were then phosphorylated using T4 polynucleotide kinase (PNK, NEB) and used in a ligation reaction with a BspQI-digested vector.

Not1 has six annotated transcript variants, with three different potential translation start sites. All but one transcript (variant F) code for proteins with the same C-terminus. The intron that gives rise to variant F shows very low inclusion in NCBI annotated RNA-seq data, so it is was assumed that most or all expressed isoforms of Not1 would harbor a tag placed at the C-terminus. This work refers to transcript variant and isoform D when using amino acid or nucleotide positions. Exon 10 was targeted for an insertion in WT Dl1 cells and primers were made to amplify a 597 bp region surrounding the stop codon. A 200 bp Ultramer containing a TEV site followed by a 3x-Flag tag was designed and ordered from IDT. Two individual homozygous knock-in clones were isolated and verified by Sanger sequencing and western blot analysis (Figure 3.20).

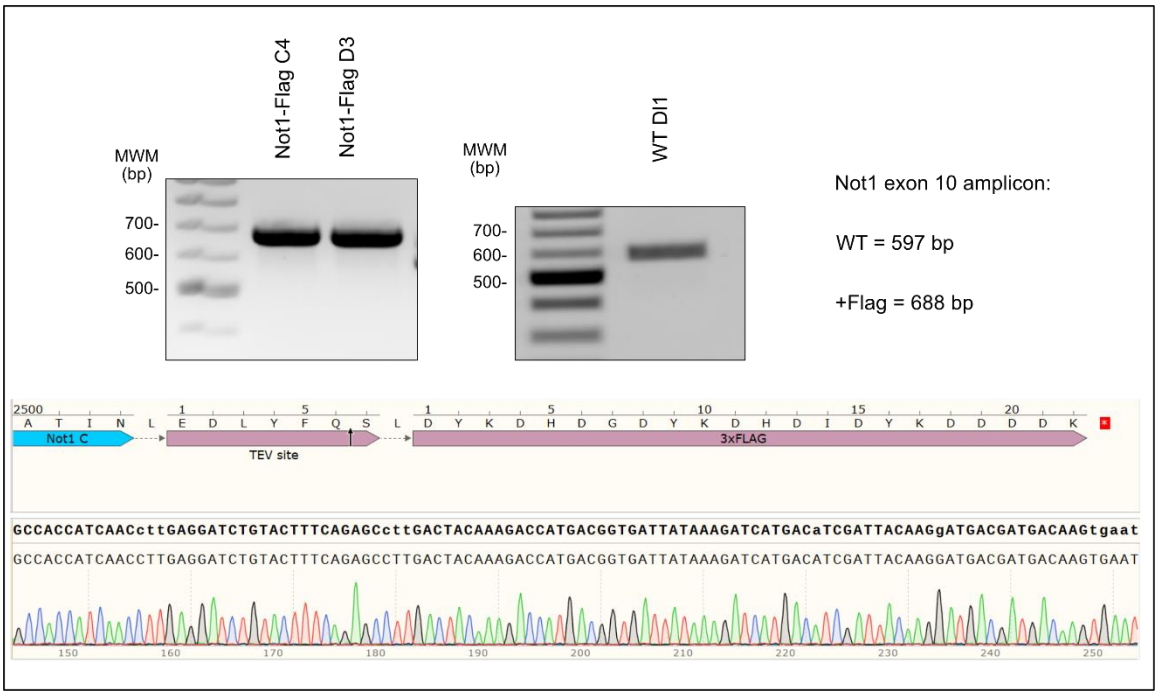


Figure 3.20: Characterization of Not1-Flag lines. (Top) DNA agarose gels showing amplicon sizes for both Not1-Flag clones (left) and WT Dl1 cells (right). (Bottom) Aligned chromatogram of sequenced amplicon for the clone used in this study (C4) using SnapGene v5.0.4. The inserted tags are shown.

Generation of CRISPR-tagged Not1 cell lines

Dl1 cells were plated at a density of 1 x 10^6 per mL in 2 mL complete Schneider's *Drosophila* Medium (SDM) in a six-well dish. Transfection mixes were set up using 4 µL FuGene HD (4:1 ratio)(Promega), 1 µg sgRNA-Cas9 plasmid DNA, 40 pmol homologous recombination (HR) template, and 100 µL serum free SDM. After a 10 minute incubation at room temperature, transfection mixes were applied dropwise to wells. One well remained untransfected to measure efficacy of Puromycin treatment. Cells were incubated at 25°C for 48 hours, then media was removed and replaced by media containing 5 µg/mL Puromycin (Gibco). After 48 to 72 hours, when untransfected cells had died, selective media was removed and replaced with complete SDM. When cells became confluent, they were passaged up to a 10-cm dish and allowed to continue growing to 80% density again. At this stage, cells were collected as a polyclonal population and approximately half were frozen and stored in liquid nitrogen. The remaining cells were diluted to isolate clones in two ways: a limiting dilution was made and cells were plated at a density of 1 to 2 cells per 100 µL into 3 to 6 96-well plates,

and 10 µL of this dilution was plated into 10-cm plates. Once visible colonies grew in the 10-cm plates, they were carefully pipetted out and transferred to single wells of a 24-well plate. For all wells of clonal isolates, once cells had achieved 80% density, they were passaged up sequentially to larger wells. Once cells had reached confluency in a 12-well plate, a portion were harvested for genomic analysis. Genomic analysis was performed by isolating genomic DNA (gDNA) from the sample using the Wizard gDNA isolation kit (Promega) and using gene-specific primers to amplify the region of intended template insertion. PCR products were analyzed on a 1% agarose gel for correct size, then purified using the Zymo DNA Clean & Concentrator 25 kit, and verified by DNA sequencing. If contamination of wild type cells was observed, cells underwent isolation procedures again until the remaining population appeared to be homozygous for the intended insertion.

Chapter 4: Endogenous targets of Pum proteins

4.1. Introduction

One of the crucial remaining questions in our understanding of Pum function is how does RD3 contribute to Pum regulation of target mRNAs? To begin to answer this question, we first needed to identify endogenous targets that we could analyze in detail in our cell system. As described in Chapter 1, several targets are well known and have important roles in development, however, little is known about target regulation in the adult cell and Pum's day-to-day role in homeostasis.

In this chapter we sought to identify *Drosophila* Pum targets that are conserved in mammals. We hypothesized that deleting RD3 or mutating F1033 and F1040 would result in a reduction of Pum's repression on these targets.

RNA immunoprecipitation assays coupled with micro-arrays (RIP-Chip) have shown that Pum binds to hundreds of transcripts in both adult *Drosophila* and embryos, and these experiments provided a starting point for identifying potential Pum targets. A 2006 RIP-Chip study expressed TAP-tagged (tandem affinity purification) Pum RBD in both *Drosophila* embryos and adult ovaries (Gerber, Luschnig et al. 2006). Over 1,000 genes were enriched in ovaries (FDR < 1%), including four well-known targets, *bicoid*, *hunchback*, *eIF4E*, and *cyclinB*. High background in the embryo dataset limited the identification to 192 genes. Further analysis of these data sets revealed only 31 genes shared between adult *Drosophila* ovaries and embryos, which could indicate distinct Pum targets at different developmental stages.

Another study in 2015 used *Drosophila* embryos to identify transcripts associated with full-length Pum using synthetic Pum antibodies (Laver, Li et al. 2015). With a 1.5-fold enrichment over negative control and FDR < 5%, 641 genes were identified as Pum targets, 300 of which were also identified in the Gerber *et al.* study in ovaries.

In addition, our laboratory performed an RNA-seq experiment in Dmel2 cells to identify targets that are increased under conditions of Pum depletion via RNAi versus LacZ (unpublished data). Although the knockdown efficiency was sub-optimal, the data provided a basis for identifying which genes could be potential targets in *Drosophila* cells.

More data is available concerning Pum targets in human or mouse cells. A 2008 RIP-chip study using human PUM1 in HeLa cells identified 726 genes as PUM1 targets (Morris, Mukherjee et al. 2008). This study also sought to compare the level of conservation of targets between yeast Puf3, *Drosophila* Pum, and human PUM1. They found that although there is no significant target conservation between Puf3 and PUM1, 502 *Drosophila* Pum targets from ovaries had human orthologs, 73 of which are also human PUM1 targets.

A similar RIP-chip experiment was performed in human HeLa S3 cells for both PUM1 and PUM2 (Galgano, Forrer et al. 2008). This study identified 1421 genes as PUM1 targets and 575 genes for PUM2, with 507 genes as targets of both Pum proteins. The authors noted that the apparent lack of identified targets for PUM2 could be due to protein degradation present in the PUM2 samples, and not necessarily indicative of a smaller binding repertoire. Comparing their results to the 2006 Gerber et al study, they found that 17% of *Drosophila* targets overlapped with human PUM1, and about 7% were shared with PUM2.

In a study designed to investigate germline homeostasis in mice, a Pum1 RIP-chip experiment was performed using mouse testes (Chen, Zheng et al. 2012). In this study, Pum1 bound to over 1500 genes with significant enrichment in gene ontology categories for p53 regulation and mitogen-activated protein kinase (MAPK) signaling. In addition, many of the target genes also showed changes in protein level when Pum1 was depleted.

In mice, Zhang *et al.* performed CLIP-seq to compare targets between neural-specific conditional Pum knock out and wild type mouse brains (Zhang, Chen et al. 2017). Their results also showed a limited pool of targets for Pum2 when compared to Pum1: Pum1 bound to 1874 genes compared to 875 for Pum2. There was significant overlap between the binding for both proteins however, with 694 targets in common.

4.2. Identification of endogenous Pum targets

Using the datasets described above, we searched for targets that could meet the following criteria: First, the targets are expressed in DL1 cells (Eric Wagner, personal communication). Second, each target has at least one PRE. Third, there is evidence of

direct binding of the target from the afore-mentioned RIP-Chip experiments. Fourth, the target has a known function. Finally, the target has a mammalian homolog. An additional consideration was whether there is a disease state associated with dysregulation of the target. From this criteria, an initial 23 targets were chosen to study, listed below in Table 4.1.

Gene	FlyBase ID	# PREs	Human homologs	binding data	disease state
Baboon	FBgn0011300	3	TGFBR1, ACVR1B+C	(1), (2), (3), (6)	yes
Ccm3	FBgn0038331	1	PDCD10	(3), (5), (6)	yes
Ches-1L	FBgn0029504	3	FOXN2+N3	(1), (3), (6)	none
Chico	FBgn0024248	1	IRS1,2,4	(3), (5), (6)	yes
CoA7	FBgn0039965	2	COA7	(1)	yes
CycK	FBgn0025674	5	CCNK	(1)	yes
E(Pc)	FBgn0000581	4	EPC1+2	(1), (3), (6)	none
Fs(1)h	FBgn0004656	4	BRD2, 3, 4, BRDT	(1), (2), (5), (6)	yes
Larp4B	FBgn0035424	4	LARP4, 4B	(1), (2), (6)	none
Msi	FBgn0011666	4	MSI1+2	(2), (4), (5), (6)	none
Myc	FBgn0262656	6	MYC	(2), (6)	yes
Nrg	FBgn0264975	1	NRCAM, L1CAM, NFASC, CHL1	(3), (5), (6)	yes
Nrv1	FBgn0015776	1	ATP1B1-4, ATP4B	(3), (5), (6)	yes
Pan	FBgn0085432	9	TCF7, TCF7L1+2, LEF1	(1), (2), (3), (4), (6)	yes
Pde11	FBgn0085370	3	PDE:2A, 4A, 4D, 5A, 6A-C, 7A, 7B, 10A, 11A	(3), (4), (5), (6)	yes
Pten	FBgn0026379	5	PTEN	(1), (2), (3), (6)	yes
Raf	FBgn0003079	2	BRAF, ARAF, RAF1	(1), (2), (3), (6)	yes
Skd	FBgn0003415	2	MED13+13L	(2), (3), (6)	yes
Spoon	FBgn0263987	3	AKAP1	(2), (3), (5), (6)	none
Strica	FBgn0033051	3	CASP3,6,7	(1), (3), (4)	none
Tis11	FBgn0011837	4	TTP (ZFP36), ZFP36L1+L2	(1), (2), (3), (6)	none
Tob	FBgn0028397	15	TOB1+2, BTG1+3	(1), (2), (3), (5), (6)	none
upSET	FBgn0036398	6	KMT2E, SETD5	(2), (6)	yes

Table 4.1: Potential endogenous targets. Binding data from (1) Gerber *et al.* 2006, (2) Laver *et al.* 2015, (3) Galgano *et al.* 2008, (4) Morris *et al.* 2008, (5) Chen *et al.* 2012, (6) Zhang *et al.* 2017.

For these experiments, we used CRISPR-Cas9 to place a Myc tag on endogenous Pum at the C-terminus. To measure the effects of Pum on the targets, we employed a knockdown strategy using RNAi. dsRNA was used to target Pum, and a non-targeting control dsRNA corresponding to *E.coli* LacZ was used for normalization. RNA was extracted from the cells and reverse-transcription and qPCR (RT-qPCR) was used to measure the target levels. Knockdown of Pum was efficient at depleting Pum protein below the level of detection for a western blot (Figure 4.1 A). Pum RNA level showed a 0.18-fold change, indicating robust depletion at the RNA level as well (Figure 4.1 B). Each experiment was performed with biological triplicates and repeated in three separate experiments. Fold changes for all tested targets are listed in Appendix E.

A 1.2-fold change cutoff ($\log(2) \approx 0.263$) was used to narrow our focus to targets with the highest biological significance. The expected level of repression seen in a minimal mRNA construct with one PRE is 1.3-fold in human cells (Bohn, Van Etten et al. 2018), however we chose to be more lenient in this definition to accommodate for other regulatory elements present in endogenous 3'-UTRs that might otherwise dampen Pum's effects. Nine targets met this criterion: Chico, Nervana1 (Nrv1), Pten, Pangolin (Pan), Phosphodiesterase 11 (Pde11), Raf, Enhancer of Polycomb (E(Pc)), Tob, and Tis11. In addition to measuring RNA levels of targets, we also measured the RNA levels of two non-targets, that is, genes that have no known PREs and are not predicted to be regulated by Pum. These non-targets served as a baseline for the effects of Pum knockdown in cells: wee augmin (wac) and Pgk. Depletion of Pum caused a decrease in the levels of non-targets, which could be interpreted as a global effect (Figure 4.1 C). However, when Pum was depleted, the levels of the PRE-containing targets increased significantly. Since these targets all have PREs and there is evidence of Pum binding to the RNA, it is probable that the observed effects were a direct result of Pum binding and initiating decay of the transcripts.

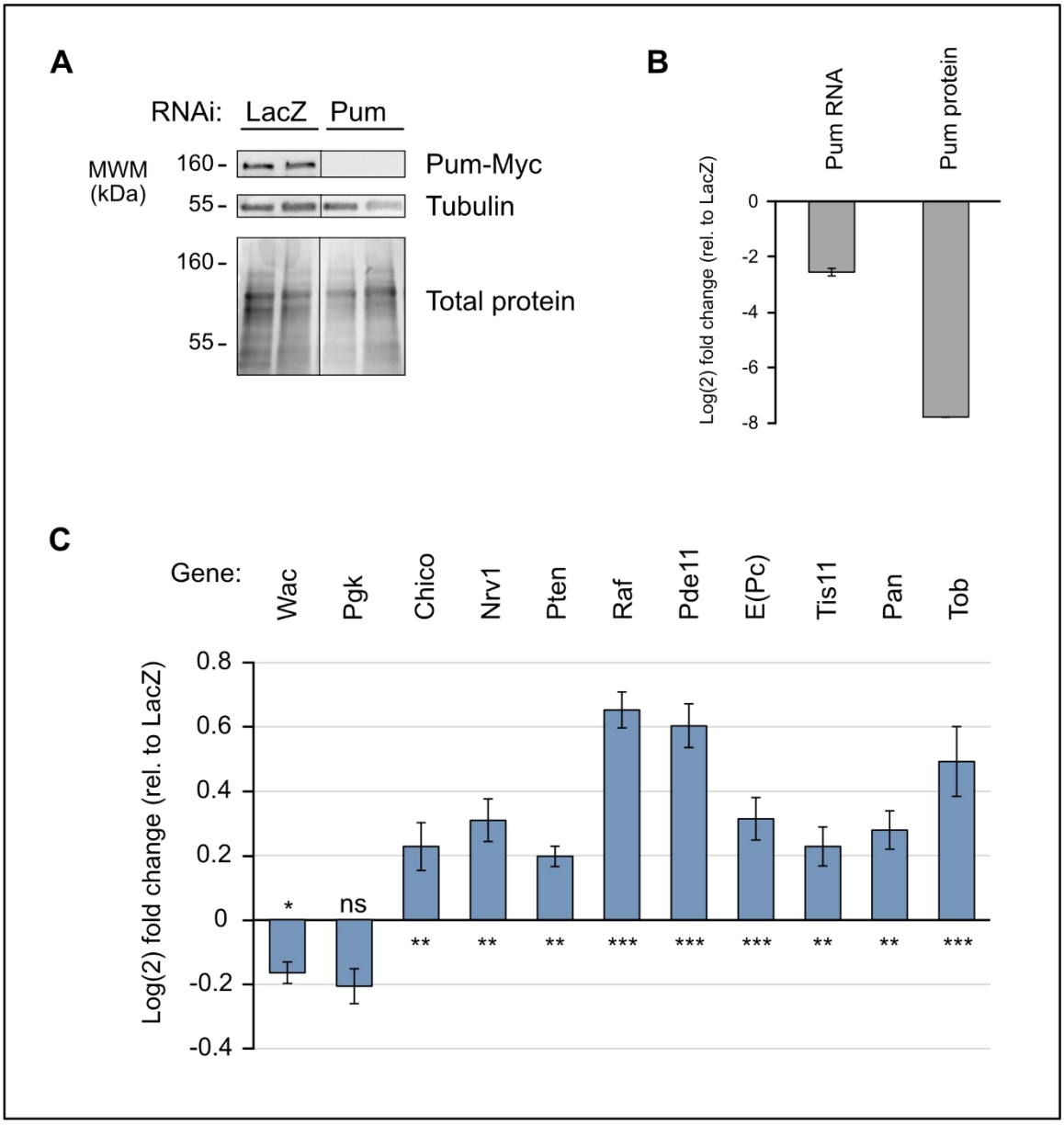


Figure 4.1: RNAi depletion of Pum increases the RNA levels of endogenous targets. (A) Western blot confirming depletion of Pum-Myc. Tubulin and total protein (Sypro Ruby) are both shown as loading controls. (B) Mean RT-qPCR measurement of Pum RNA and quantification of Pum protein level demonstrating efficient depletion of Pum. (C) Mean target RNA levels as measured by RT-qPCR under depletion of Pum using RNAi (n = 9). RT-qPCR experiments are normalized to ribosomal protein RPL32 and Pum RNAi compared to LacZ. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Error bars represent standard error of the mean (SEM).

4.3. Validation of Pum targets

To demonstrate that Pum regulation of these targets is dependent on the PRE, we cloned the 3'-UTRs of Raf, Pde11, Nrv1, and Chico into luciferase reporters. Mutant reporters of each target were made by mutating each PRE sequence from UGUANAUA to ACAANAUA (Figure 4.2). These experiments were performed in Dl1 cells, using endogenous Pum to measure the effect of mutating the PRE on the target reporters. As a positive control, we included a PRE reporter consisting of 3 PREs in a minimal 3'-UTR (3x min). Each reporter was normalized to its respective mutant. Experiments were carried out in biological quadruplicates and repeated three times. Pum repression of the target reporters was dependent on the presence of a functional PRE (Figure 4.2).

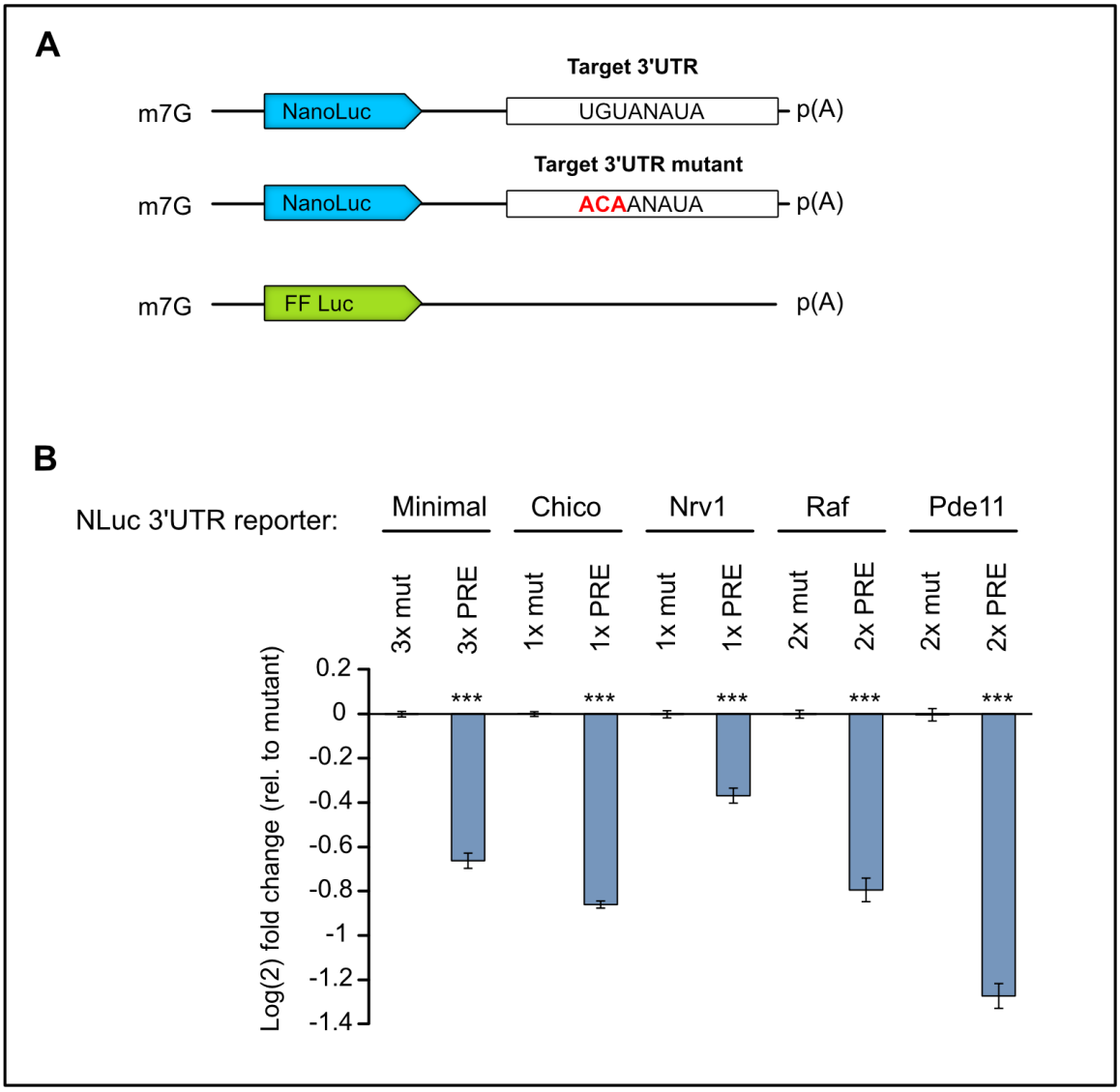


Figure 4.2: Pum repression of target reporters is PRE-dependent. (A) Representation of target reporters. Each 3'-UTR was cloned into a NanoLuc reporter with WT sequence or mutated PREs. FFLuc is an internal control. (B) Mean log(2) fold change of target reporters showing SEM (n = 12). Each WT 3'-UTR (x PRE) is normalized to the respective mutant PRE (x mut). Significance against mutant reporters is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

4.4. Contribution of RD3 to repression of endogenous targets

Having established PRE-dependent repression of our target reporters, we next asked what contribution RD3 has on regulating the target reporters. Previous

experiments in Dmel-2 cells had shown that redundancy among the Pum RDs can compensate for the deletion of a single RD (Weidmann and Goldstrohm 2012), however these experiments were performed using minimal 3'-UTR reporters.

To test this, we overexpressed full-length Pum constructs in DL1 cells or Pum with RD3 removed (Pum ΔRD3). These effectors were normalized to overexpressed eGFP. As a negative control, we included a Pum R7 mutant. This mutant has amino acid substitutions in repeat 7 of the RBD (S1342A, N1343A, E1346A), rendering it incapable of binding RNA (Weidmann and Goldstrohm 2012). As such, its effects on the target reporters should be minimal. Experiments were carried out in biological quadruplicates and repeated three times. Overexpression of full-length Pum or Pum ΔRD3 effectively repressed the reporters over the endogenous amount of Pum, but the R7 mutant had minimal effect (Figure 4.3).

Additionally, the effects of deleting RD3 were different for certain targets. There was no significant difference between Pum and Pum ΔRD3 on the Chico reporter, and a small difference on the Nrv1 reporter (Figure 4.3 A, B). Removing RD3 appeared to enhance repression on the Pde11 target reporter (+7%), while it was detrimental on the Raf reporter (-8%) (Figure 4.3 C, D).

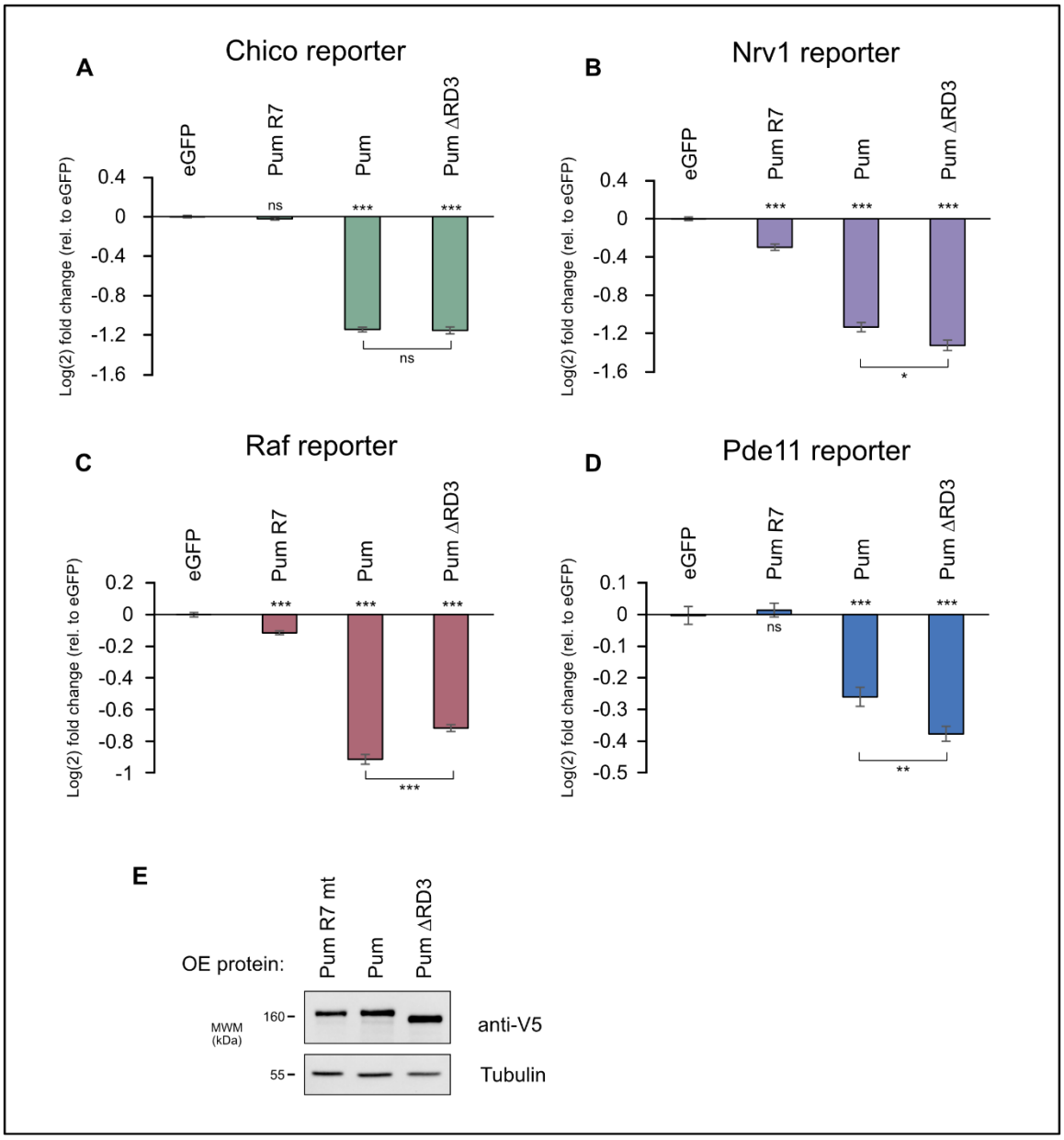


Figure 4.3: Overexpression of Pum represses target reporters. (A-D) Mean log(2) fold change relative to eGFP is shown for all reporters with SEM (n = 12). Significance against eGFP is shown along the x axis. Significance between Pum and Pum ΔRD3 is shown below the graph. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (E) Western blot showing expression of V5-tagged effectors with tubulin as a loading control.

To further explore the contribution of RD3 to repression and the effect of mutating F1033 and F1040, we switched to using a Pum-null cell line that we created using CRISPR-Cas9 gene editing. Since the RBD is necessary for binding of Pum to the

target reporter, we overexpressed either the RBD alone, or the contiguous region RD3-RBD with its mutant, RD3-RBD F1033A+F1040A (RD3-RBD FF). As a positive control for the luciferase assays, we included a reporter with a minimal 3'-UTR, consisting simply of 3 PRE sites (3xPRE), and its corresponding mutant, 3xACA. The Pum RBD has a small amount of activity when expressed alone (Weidmann and Goldstrohm 2012), so we normalized these experiments to the RNA binding R7 mutant RBD (RBD R7). Experiments were carried out in biological quadruplicates and repeated three times. RD3 enhanced repression over the RBD alone, but this additional repressive activity was abrogated when both phenylalanines were mutated (Figure 4.4). The effects of the FF mutation on the Nrv1 and 3x min reporters were especially striking. The RD3-RBD FF mutation caused a loss of activity equivalent to that of the R7 mutant, which cannot bind RNA.

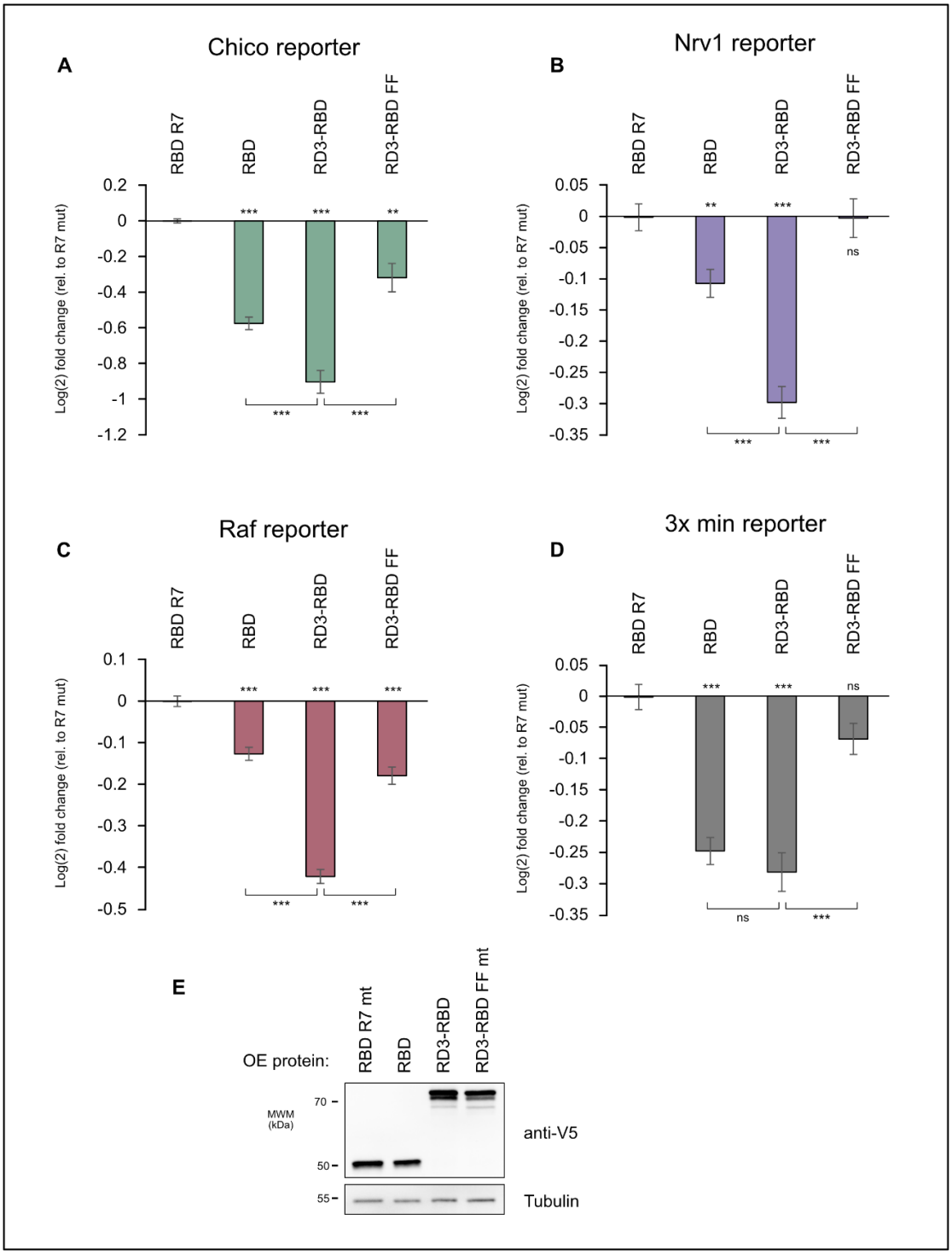


Figure 4.4: Overexpression of RD3 represses target reporters. (A-D) Mean log(2) fold change relative to RBD R7 mutant is shown for all reporters with SEM (n = 12). Significance against RBD R7 is shown along the x axis. Significance between RBD, RD3-RBD, and RD3-RBD FF is shown below the graph. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. (E) Western blot showing expression of V5-tagged effectors with tubulin as a loading control.

4.5. Pum depletion causes an increase in target protein levels

RNA level does not always correlate with protein level. To demonstrate that the effects of Pum depletion on the targets were biologically relevant, we asked if we could also detect a change in the protein levels of targets when Pum was depleted. This is especially important for signaling molecules and other proteins that the cell keeps under tight regulation. The targets Raf and Pde11 are two such proteins.

Using CRISPR-Cas9 editing, we placed V5 tags on the endogenous Raf and Pde11 proteins in our Pum-Myc cell line. RNAi was used to knock down Pum in these cells and quantitative western blotting was used to measure the resulting protein levels of the targets. Each knockdown experiment was performed in triplicate and repeated three times for reproducibility. The average change in protein level for both Pde11 and Raf was then assessed and compared to the change in RNA level (Figure 4.5 A). For Raf and Pde11, a change in RNA level corresponded to a change in the protein level.

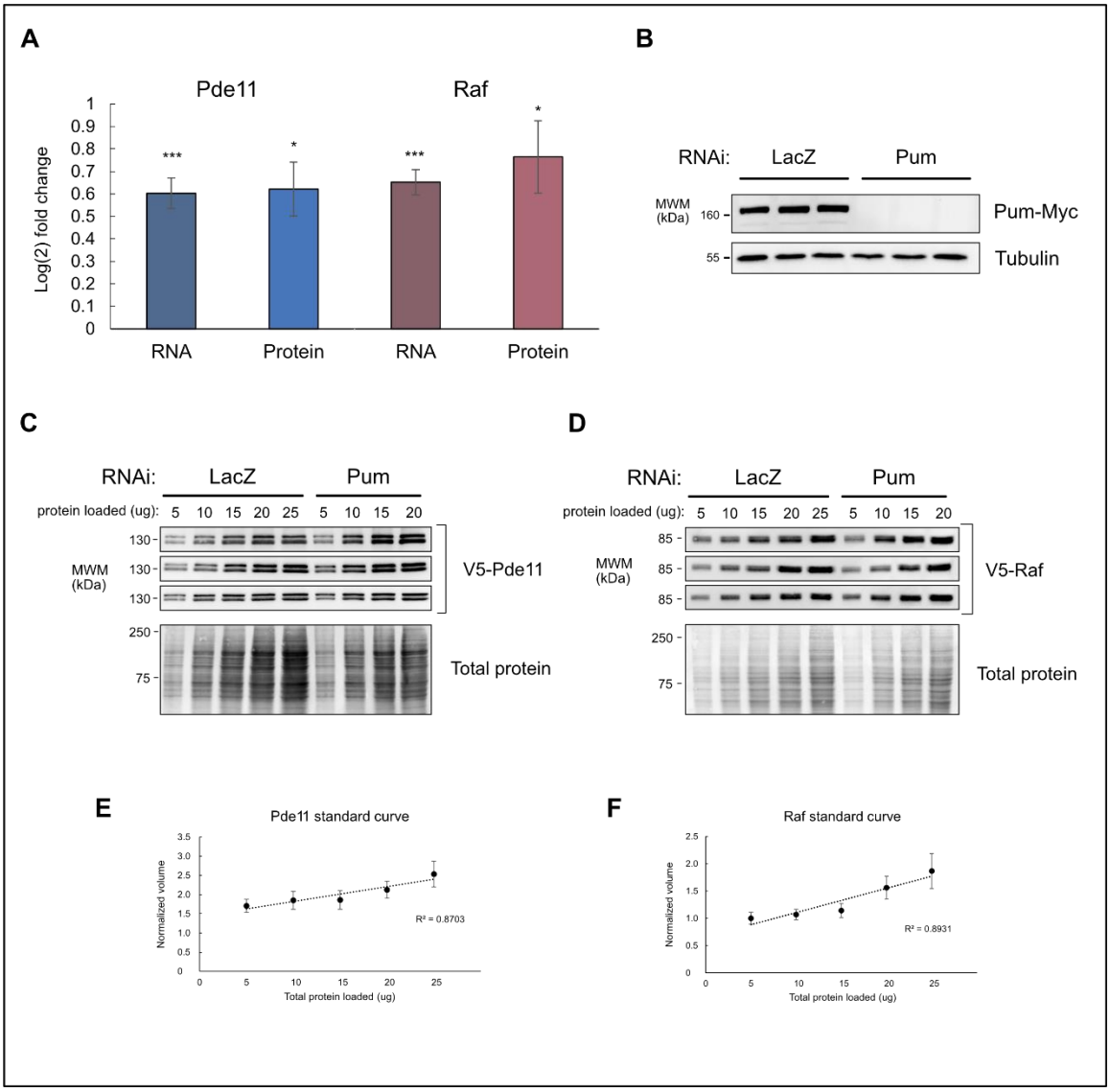


Figure 4.5: Pum depletion causes an increase in target protein levels. Results of quantitative western blotting for V5-Raf and V5-Pde11. (A) Mean log(2) fold change for RNA levels as measured previously by RT-qPCR, and protein levels. Error bars represent standard error of the mean. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. n = 9 for RNA data; n = 9 for protein data. (B) Representative western blot showing depletion of Myc-tagged Pum with Tubulin as a loading control. (C, D) Representative western blots of V5-Pde11 (C) and V5-Raf (D) with total protein staining by Sypro Ruby. Titrations of total protein are loaded for the LacZ NTC condition and Pum RNAi condition. Band density is normalized to total protein as described in Methods. (E, F) Standard curve for Pde11 (E) and Raf (F). Normalized band volume for the LacZ condition is plotted against total protein loaded (as measured by Lowry DC assay). n = 9 for each data point.

4.6. Not1 as the effector of Pum-mediated decay of endogenous targets

Having established that Pum uses the CNOT complex for repression, we next asked if our Pum targets were also affected similarly when Not1 is depleted. Although we expected global changes when knocking down a major deadenylase, not all mRNAs are regulated by the CNOT complex. Since there is no evidence of Pum interacting with the Pan complex, we expected that all of Pum's targets would be specifically affected by Not1 depletion.

We tested the RNA levels of the nine previous targets by using RNAi to knock down Not1 in our Not1-Flag Dl1 cells. Target levels were measured by RT-qPCR and compared to a LacZ non-targeting control. Most of the tested Pum targets were affected by Not1 knockdown (Figure 4.6).

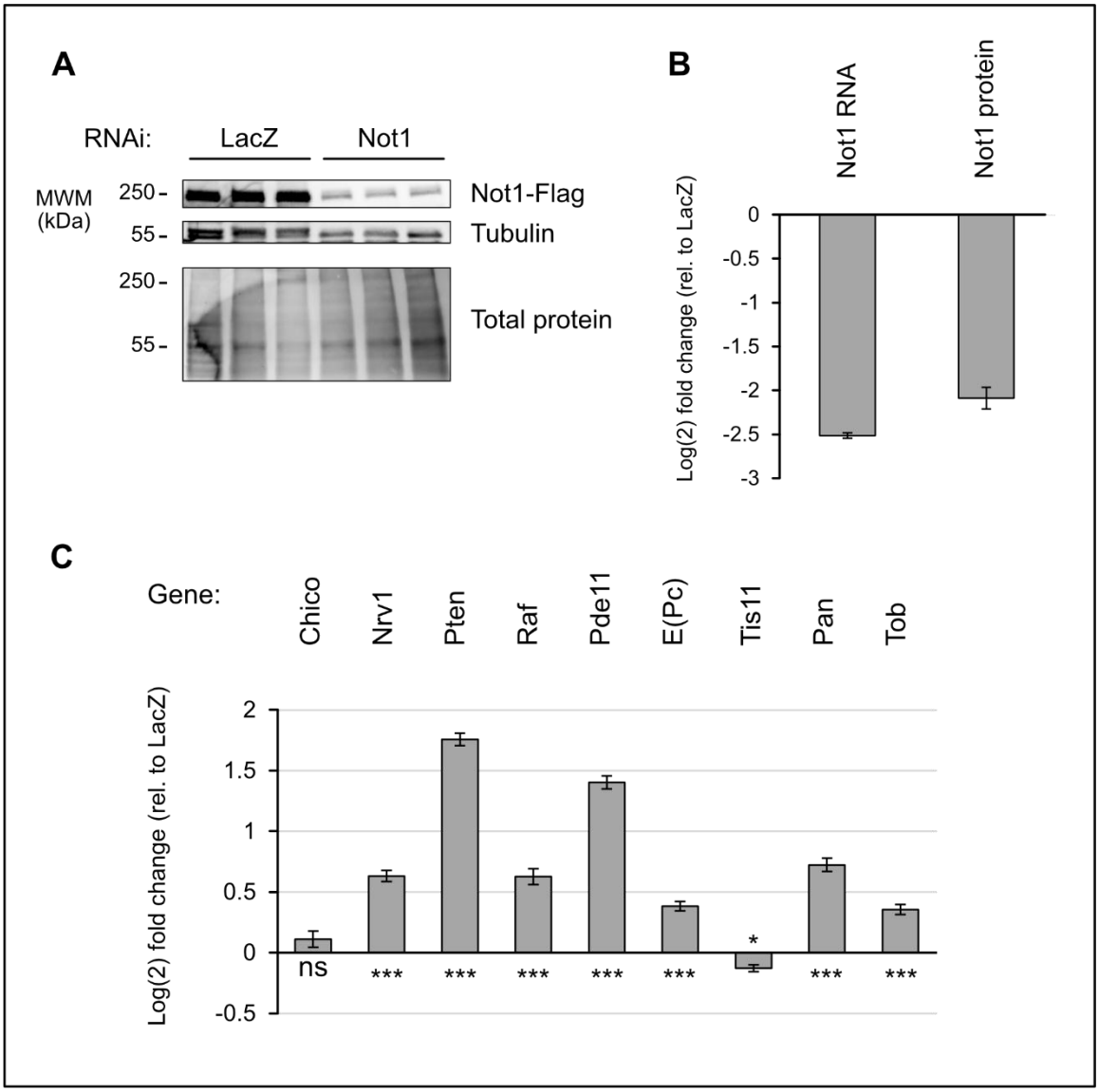


Figure 4.6: RNAi depletion of Not1 increases the RNA levels of endogenous Pum targets. (A) Western blot confirming depletion of Not1-Flag. Tubulin and total protein (Sypro Ruby) are both shown as loading controls. (B) Mean RT-qPCR measurement of Not1 RNA and quantification of Not1 protein level demonstrating depletion of Not1. (C) Mean target RNA levels as measured by RT-qPCR under depletion of Not1 using RNAi (n = 9). RT-qPCR experiments are normalized to Act5c and Not1 RNAi compared to LacZ. Significance calling is as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Error bars represent standard error of the mean (SEM).

4.7. Summary and discussion

Much of our understanding of Pum targets comes from developmental or tissue specific analysis. Our experiments demonstrate the importance of Pum-mediated decay outside of these contexts and show that regulation by Pum is relevant in a cell-based system. Furthermore, RD3 contributes to the repression of endogenous targets.

We identified at least nine endogenous targets of Pum and showed that their RNA levels were affected by Pum depletion. We further verified Pum-mediated repression by demonstrating that regulation of four of these targets (Pde11, Raf, Nrv1, and Chico) was dependent on the Pum binding sites present in the 3'-UTRs. Additionally, overexpression of Pum caused a decrease in the target reporter levels. We also showed that RD3 can repress these target reporters, and that mutating the phenylalanine residues F1033 and F1040 severely impacted the repressive ability of RD3. We also showed that depletion of Not1 affected Pum targets, as the effector of Pum-mediated decay. Finally, we demonstrated that in addition to affecting RNA levels of endogenous targets, depletion of Pum caused an increase in the protein levels of both Raf and Pde11.

One of the observations worth noting is the fact that the number of PREs in a target's 3'-UTR does not correlate to repression level by Pum. Presumably, the presence of a single PRE is enough to enact decay by Pum, and the presence of multiple PREs serves only to increase the odds that Pum will bind to the transcript (Bohn, Van Etten et al. 2018, Wolfe, Schagat et al. 2020). Even so, several of the genes we tested did not change appreciably under Pum depletion regardless of PRE number. There are several explanations for this. One is the presence of other *cis* regulatory elements in the 3'-UTR. Pum is but one of many regulatory proteins bound to a target RNA at any given moment. Other RBPs might sterically block the PRE and prevent Pum binding or force the transcript toward a different fate. Pum might also be bound to the PRE, but otherwise be prevented from recruiting the CNOT complex by associations with other proteins, or by autoregulation. A recent study of transcripts regulated by human PUM1 and PUM2 identified an enriched motif present in genes that were downregulated when PUM1/2 were depleted (Bohn, Van Etten et al. 2018). This could indicate a mode of combinatorial control or the unknown protein(s) that bind to this motif may supersede Pum's activity.

The position of the PRE could also be a factor. The most active PREs tend to be located in the 3'-UTR versus the CDS or 5'-UTR (Bohn, Van Etten et al. 2018, Wolfe, Schagat et al. 2020). The reasons for this are unclear, but given the known mechanisms of Pum action involving recruiting deadenylases and inhibiting PABP (Van Etten, Schagat et al. 2012), proximity to the poly(A) tail could be a contributing factor.

The inherent structure of the RNA itself might inhibit Pum binding. The common depiction of RNA as a linear molecule is grossly misleading, as RNA forms many structures *in vivo*. These include *cis* base pairing, G-quadruplex structures, pseudoknots, hairpins and helices (Ganser, Kelly et al. 2019). Any of these structures can form to occlude an RBP binding site.

The dynamic interplay of RNA regulatory factors can be seen in the target 3'-UTR reporter assays for Pde11, Raf, Chico, and Nrv1. All other conditions being equal, each reporter had a different response to Pum independent of PRE number (Figure 4.2). Additionally, overexpressing Pum or Pum ΔRD3 affected the reporters differently. For the Pde11 reporter, the loss of RD3 resulted in an increase in repression by Pum, while the opposite was true for the Raf reporter (Figure 4.3). Furthermore, the effects of mutating the phenylalanine residues in RD3 were much more pronounced for the Nrv1 reporter (Figure 4.4).

One of the top targets identified in our RNAi screen was Tob, a target that has 15 PREs. Tob has a very long 3'-UTR (5,000+ bp) and we were not able to clone it fully into our reporter assay system. Instead, we cloned an area that covered 5 of the PRE sites. However, this reporter did not show PRE-dependent repression, nor did a reporter made from the Ccm3 3'-UTR with 2 PREs (data not shown). It is unknown what other factors might contribute to regulation of these targets or whether Pum is acting on them at all.

The data presented here illustrates the complexity surrounding the modality of Pum repression. The larger question of what makes a PRE functional still remains a mystery and a topic of great interest.

4.7.1. Biological significance of targets

The Pum targets identified in this chapter have many different cellular functions. This set of Pum targets, while just a small sampling of the potential hundreds of genes regulated by Pum, demonstrates the importance and scope of the regulatory power of

Pum proteins. The Pum targets described here function in multiple cell-signaling pathways, regulate development, and control cell growth and tumor progression. Here we summarize what is known about these targets in *Drosophila* and mammals.

Cyclic GMP and AMP (cAMP and cGMP) are important cytoplasmic signaling molecules that regulate diverse cellular functions. Levels of cAMP and cGMP are tightly controlled by their interconversion from linear and cyclic forms by adenyl or guanyl cyclases and phosphodiesterases (PDEs). PDEs are responsible for catalyzing the conversion from cGMP and cAMP to 5' GMP and 5' AMP. In humans, there are ten families of PDEs, with most members exhibiting distinct tissue localization, substrate specificity, allosteric activators and regulation. Because of their roles in signal transduction pathways, PDEs are important drug targets. Pde11 is one of six PDEs in *Drosophila*, exhibiting dual specificity for both cGMP and cAMP. Of the fly PDEs, Pde11 shows the highest affinity for cGMP (Day, Dow et al. 2005).

Pde11 has been best studied in human and mouse systems, where PDE11 is important for male testicular function and controlling cAMP levels (Wayman, Phillips et al. 2005, Levy, Szarek et al. 2021). PDEs are often implicated in adrenal tumors, where levels of cAMP are dysregulated (Szarek and Stratakis 2014). In humans, mutations of Pde11 were discovered in patients with idiopathic Cushing syndrome. These mutations resulted in decreased protein levels of PDE11 in adrenal tumor tissues, increased levels of cyclic nucleotides, and increased phosphorylation of cAMP-responsive element binding protein (CREB) (Horvath, Boikos et al. 2006). In mouse brain, Pde11a KO also affects societal learning and memory formation (Makhlouf, Kshirsagar et al. 2006, Kelly, Logue et al. 2010, Kelly 2017).

Drosophila Pde11 shares a 77% sequence identity with human PDE11 in the catalytic domain, and 38% in the rest of the protein (Day, Dow et al. 2005). In *Drosophila*, the Pde11 gene has six annotated transcripts with four unique 3'-UTRs. A portion of the 3'-UTR common to all transcripts contains 2 perfect PREs and an additional partial (7 nt) PRE. Importantly, these PREs are conserved in human PDE5A and PDE11A, the two closest related human PDEs.

Raf and its human homologs BRAF, ARAF, and Raf1 are serine/threonine kinases that act in the Tor, Egfr and sevenless signaling cascades. These pathways are crucial for cell growth and proliferation. As the activity of these kinases lead to

amplification of downstream events, dysregulation of Raf proteins often lead to cancer. As such, their roles as oncogenes have been studied for decades (Nishida, Hata et al. 1988, Brose, Volpe et al. 2002, Maurer, Tarkowski et al. 2011).

In *Drosophila*, Raf also has roles in memory, where its overexpression leads to stabilization of labile memory (Zhang, Li et al. 2018). A large-scale TAP-MS (tandem affinity purification-mass spectrometry) study uncovered hundreds of protein interactors of *Drosophila* Raf, highlighting the interconnectivity of this protein in the MAPK signaling cascade (Friedman, Tucker et al. 2011).

Drosophila Raf shares 46% amino acid identity to human RAF1, 44% to ARAF, and 43% identity to BRAF (DRSC). The *Drosophila* Raf gene encodes two transcript variants, both of which translate into a single protein isoform. Both transcripts contain 2 perfect PREs in the 3' UTR, and the longest transcript contains an additional partial (7 nt) PRE in the 5' UTR. The method of post-transcriptional control by Pums is likely conserved in humans as well, with all three human Raf homologs (ARAF, BRAF, RAF1) possessing at least one PRE in their 3'-UTRs. The human ARAF transcript has also been identified in a genome-wide RIP-Chip for binding to both PUM1 and PUM2, and the mouse homolog is also bound by Pum (Galgano, Forrer et al. 2008, Zhang, Chen et al. 2017). Additionally, human RAF1 was shown to functionally compensate for *Drosophila* Raf in mutant flies (Baek, Fabian et al. 1996).

Chico, first identified in a screen for mutations causing small body size in *Drosophila*, is an insulin receptor substrate, homologous to human IRS1,2, and IRS4. In addition to controlling body size by reducing both cell size and number, chico mutants also produced nearly double the amount of lipids as WT flies (Böhni, Riesgo-Escovar et al. 1999). The human insulin receptor substrate IRS1 contains 8 PREs in its 3'-UTR.

Nrv1 (Nervana1) is one of three beta subunits of the sodium-potassium ion channel pump (Na+ K+ ATPase). It is orthologous to human beta subunits ATP1B1-B4, and ATP4B. The Na+ K+ ATPase ion pumps are essential for maintaining cell polarity, which allows for the uptake of molecules and establishes resting membrane potential. Nrv1 is expressed in muscle tissues and non-neuronal tissues (Xu, Sun et al. 1999). Each of the beta subunits exhibit distinct tissue localization indicating unique functions, however the specific roles of Nrv1 in *Drosophila* have not been well studied (Paul, Palladino et al. 2007).

Tis11 (TTP, Tristetraprolin in humans) is a conserved RBP that regulates mRNAs in a manner similar to Pum. Tis11 binds to AU-rich elements (AREs) in the 3' UTR of target mRNAs via tandem zinc fingers and recruits the CNOT deadenylase complex (Lai, Carballo et al. 2000). Tis11 mRNA expression is induced following mitogenic stimulation (Sanduja, Blanco et al. 2012).

Pten (phosphatase and tensin homolog) is a dual specificity phosphatase with activity for phosphoinositide lipids and tyrosine or serine/threonine phosphorylated residues. Pten inhibits signaling from phosphoinositide 3-kinase (PI3K) by dephosphorylating phosphatidylinositol 3,4,5-trisphosphate and negatively regulating the insulin receptor pathway (Goberdhan, Paricio et al. 1999). Both *Drosophila* Pten and its mammalian homolog PTEN have been extensively studied for their roles as tumor suppressors in regulating cell growth (Shi, Paluch et al. 2012, Texada, Koyama et al. 2020).

Pan (Pangolin, also known as dTCF) is a transcription factor in the Wnt signaling pathway that acts as a repressor or an activator. Wnt signaling is important for cell-to-cell communication during development and to maintain cellular homeostasis. Upon activation of the pathway, intracellular levels of Armadillo (β-catenin) protein rise and bind to Pan, switching Pan from a transcriptional repressor to an activator. Pan thus facilitates the transcription of genes in response to the received Wnt signal (Fiedler, Graeb et al. 2015).

Tob (transducer of ERBB2) is most closely related to mammalian TOB1. The Tob family of proteins are well known anti-proliferative proteins which also act by recruiting the CNOT complex via interactions with Pop2 (human CNOT7/8) (Winkler 2010, Doidge, Mittal et al. 2012). Having no intrinsic RNA binding ability of their own, Tob proteins target other RBPs like CPEB or PABP to illicit decay of the transcripts they are bound to (Hosoda, Funakoshi et al. 2011).

Enhancer of polycomb (E(Pc)) was first identified for its ability to suppress position-effect variegation in *Drosophila* (Sato 1983, Elgin and Reuter 2013). It is one of the non-catalytic subunits of the NuA4 histone acetyltransferase complex (Doyon and Cote 2004). More recently, roles for E(Pc) have been found in hematopoiesis (Owusu-Ansah and Banerjee 2009), cell cycle control, DNA repair, and cell differentiation (Searle and Pillus 2018).

4.8. Materials and methods

Cell culture

Drosophila Dmel2 cells were cultured in Gibco SF-900III media with no additional supplements. Cells were passaged every 3 to 4 days and discarded once they reached passage 30. *Drosophila* Dl1 cells were a gift from Eric Wagner. Dl1 cells were cultured in Schneider's *Drosophila* media (SDM, Gibco) containing additional glutamine (1x GlutaMAX, Gibco), 1x anti-microbial/anti-fungal (Anti/anti, Thermo Fisher), and 10% heat inactivated fetal bovine serum (FBS, GenClone). FBS was heat inactivated by incubating at 56 to 60°C for 30 minutes. DL1 cells were passaged every 3 to 4 days and discarded after passage number 40. All *Drosophila* cells lines were cultured at 25°C in a temperature and humidity-controlled incubator. DL1 cells were frozen from 80% density in 10-cm dishes, collected and centrifuged in 15 mL tubes at 800 x g for 3 minutes. 2 mL of used media was reserved; 500 µL DMSO and 2.5 mL of complete SDM were added to the tube and cells were resuspended. Cells were aliquoted into 3 cryogenic storage tubes, placed in a Styrofoam tray and stored at -80°C for 2 to 7 days before transfer to liquid nitrogen.

Luciferase assays

100 µL of *Drosophila* cells were seeded in 96-well plates at a density of 1.5 x 10^6 per mL. Four wells in each condition served as biological replicates for luciferase measurements, while three wells in each condition were reserved for protein analysis. Transfection master mixes were set up using 1 ng of Firefly Luciferase (FF) per well (5 ng for Dmel2), 1 ng of Nano Luciferase (NLuc) reporter (5 ng for Dmel2), and 98 ng of either effector plasmid or balanced with empty pIZ plasmid (90 ng for Dmel2). FuGene HD transfection reagent (Promega) was used at a 4:1 ratio (0.4 µL/well) and serum free media was added to master mixes to bring the level of master mix transfected per well to 10 µL. Master mixes were incubated for 5 to 10 minutes before adding to wells of plated cells. Plates were incubated for 48 to 72 hours before harvest and analysis. For analysis, 100 µL of OneGlo Firefly reagent (Promega) was added to each well, mixed and incubated at room temperature for 10 minutes. Luminescence was measured on a Glo-Max instrument under default settings. 100 µL of NanoGlo Stop-n-Glo reagent

(Promega) was then added to each well, mixed and incubated at room temperature for 15 minutes before reading luminescence.

RNAi

DNA templates for *in vitro* transcription were made by amplifying from either plasmid DNA or cDNA using primers containing T7 RNA polymerase promotor sites. Templates for Pum knockdown target a 597 bp region in the RBD and Not1 templates target a 536 bp region in the Caf40 binding domain. DNA templates were purified using either Qiagen's QIAquick PCR purification kit or Zymo DNA Clean & Concentrator kit. dsRNA was made using 1 µg of template with NEB's HiScribe T7 RNA polymerase. dsRNA was purified using either Zymo RNA Clean & Concentrator or Monarch RNA Cleanup kit (NEB).

Drosophila Dl1 cells were plated in 6-well plates at a density of 1 x 10^6 cells/ mL for 4-day knockdowns, or 1.5 x 10^6 cells/mL for 3-day knockdowns. Cells were incubated in 10 to 20 µg (20 µg for Not1) dsRNA for 60 minutes in 1 mL serum-free media. 2 mL complete media was added after dsRNA bathing step and cells were incubated at 25°C for 68 to 90 hours. At harvest, the media was removed and the cells were resuspended in 2 mL PBS. 1.5 mL of cells were taken for RNA, and the remaining 500 µL were processed for protein expression.

Western blotting

For luciferase assays, three wells of each condition were harvested by pipette-mixing and transferring cells to 1.5 mL Eppendorf tubes. Cells were centrifuged at 900 x g for 4 minutes, media was removed, and the cells were resuspended in 40 µL RIPA buffer with additional 2x protease inhibitors (Roche, Sigma). For verification of knockdown in tagged cell lines, harvested cells were resuspended in 2 mL PBS. 1.5 mL of cells were taken for RNA, and the remaining 500 µL were pelleted and resuspended in 75 µL RIPA buffer with 2x protease inhibitors (Roche). Further processing was the same. After a 10 minute incubation on ice, lysates were collected by centrifuging at 21,000 x g for 10 minutes and removing the supernatant. The protein concentration was measured using the Bio-Rad DC protein assay kit according to the manufacturer's directions.

5 µg (from luciferase assays) or 20 µg (for Pum or Not1 knockdown verification) of total protein extract was incubated with an equal volume of 2x SDS loading buffer at 85°C for ten minutes. Samples were loaded on 12% SDS-PAGE gels (either laboratory-made or purchased precast from BioRad) and run at 120 to 180 volts until dye front reached the bottom of the gel. Gels were transferred to PVDF membranes (Millipore) for either 1.5 hours at 65 volts or overnight at 35 volts. Following transfer, membranes were either blocked in blotto for 1 hour or allowed to dry completely, then primary antibodies were applied for one hour at room temperature, or incubated overnight at 4°C. Membranes were washed three times for 10 minutes in blotto, then secondary antibodies were applied for 1 hour at room temperature. Membranes were washed three times for ten minutes each in blotto, then rinsed with 1x PBS before applying chemiluminescent substrate (Pierce- Thermo Fisher, or Immobilon- Millipore) Blots were imaged using the auto exposure setting on a ChemiDoc Touch (BioRad).

Data analysis

Luciferase assays: Samples were normalized for potential variation in transfection efficiency by calculating the ratio of NLuc to FF. The repression activity of each effector was calculated as fold change by dividing each sample well NLuc:FF ratio by the average NLuc:FF ratio of the control condition (eGFP, R7 RNA binding mutant, or ACA mutant reporter.) Repression activity was then graphed as log(2) fold change. Statistics for luciferase assays were performed using Student's unpaired T-test with either equal or unequal variance as dictated by F-test results. Significance calling was based on the convention $p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***. Each experiment consisted of 3 to 4 biological replicates and each experiment was repeated at least 3 individual times.

RT-qPCR: Three biological replicates for NTC or WT and 3 biological replicates for experimental condition were used for each experiment. Each experimental condition was compared to each NTC/WT condition for 3 measurements per biological sample, and 9 total measurements per experiment. Fold change for each experiment was determined using the Pfaffl method where fold change is represented by the equation $FC = \frac{(E_{Target})^{\Delta Ct_{Target}}}{(E_{Ref})^{\Delta Ct_{Ref}}}$. E is defined as primer efficiency for each gene, and ∆Ct is Ct of

the calibrator condition (LacZ) minus the Ct of the test condition (Pum KD or Not1KD.) Reference genes were RPL32 for Pum KD and Act5c for Not1 KD. Unpaired Student's T tests with variances determined by F tests were performed using raw ΔCt values for each gene. Each experiment was repeated 3 times, for a total of 27 measurements.

RT-qPCR

RNA extraction was performed using the Maxwell RSC simply RNA Tissue RNA extraction kit and concentration was determined using the NanoDrop spectrophotometer. A total of 10 µg RNA was taken from each sample for reverse transcription (RT) using GoScript reverse transcriptase (5 µg for RT, 5 µg for noRT samples.) RT reactions were primed with random hexamers and carried out according to the manufacturer's protocol. cDNA was then diluted with 100 µL water to approximately 41 ng/µL. Quantitative PCR (qPCR) was set up using Promega's Go-Taq qPCR Master Mix with 2 µL of each cDNA/noRT sample in a 20 µL reaction volume using either 100 nM or 200 nM final concentration qPCR primers for each gene. Reactions were performed on the Bio-Rad CFX96 instrument using the following cycling parameters: 3 min 98°C, [10 sec 95°C, 30 sec 62/63/64°C, 40 sec 72°C + image] x 39, 60°C-90°C melt curve + image. Individual annealing temps for each primer set are listed in Appendix D.

Cloning target reporters

Endogenous target reporters were created by ordering the 3'-UTR of each transcript as a gBlock from IDT to insert into the pAc 5.4 NLuc vector using XhoI and NotI restriction sites. For Pde11, nucleotides 5762 to 6368 of transcript variant E were used. This region of the 3'-UTR is common to all transcript variants and contains 2 perfect PREs.

Plasmids

Archive number	**Plasmid name**	**Creator**
ACG864	pIZ Dm Pumilio V5 His6	Aaron Goldstrohm
CAW003	pIZ Dm Pumilio R7mut SNVVE to AAVVA	Chase Weidmann
ACG866	pIZ Dm Pumilio RBD	Aaron Goldstrohm
CAW391	pIZ PumΔRD3 V5H6	Chase Weidmann

Archive number	Plasmid name	Creator
ACG544	pIZ GFP	Marzluf Lab
RMA0008	pIZ Dm Pum RD3-RBD	Rene Arvola
RJH229	pIZ RD3-RBD FF Mut	Rebecca Haugen
CAW004	pIZ Dm Pum RBD R7mut SNVVE to AAVVA	Chase Weidmann
RJH230	pAc 5.4 NLuc 3xPRE (19 nt)	Rebecca Haugen
RJH231	pAc 5.4 NLuc 3xPRE ACA mt (19 nt)	Rebecca Haugen
CAW061	pIZ MS2CP-GFP V5H6	Chase Weidmann
RJH232	pAc 5.4 NLuc Chico 3'-UTR	Rebecca Haugen
RJH233	pAc 5.4 NLuc Chico 3'-UTR 2x ACA mt	Rebecca Haugen
RJH234	pAc 5.4 NLuc Nrv1 3'-UTR	Rebecca Haugen
RJH235	pAc 5.4 NLuc Nrv1 3'-UTR 1x ACA mt	Rebecca Haugen
RJH236	pAc 5.4 NLuc Pde11 3'-UTR	Rebecca Haugen
RJH237	pAc 5.4 NLuc Pde11 3'-UTR 2xACA mt	Rebecca Haugen
RJH238	pAc 5.4 NLuc Raf 3'-UTR	Rebecca Haugen
RJH239	pAc 5.4 NLuc Raf 3'-UTR 2xACA mt	Rebecca Haugen

Table 4.2: Plasmids used in Chapter 4. A full list of plasmids can be found in Appendix C.

Antibodies

All antibodies used in this study can be found listed in Appendix B. Mouse anti-V5 primary antibody (Invitrogen, Cat# R960-25) was diluted 1:5000 in blotto. Mouse anti-Tubulin (CST, Cat# 3873) was diluted 1:1000 in blotto. Rabbit anti-Myc primary antibody (CST Cat# 2278S) was diluted 1:5000 in TBST. Mouse anti-Flag (Sigma Cat# F3165-1MG) was diluted 1:5000 in blotto. Goat anti-rabbit-HRP secondary antibody (CST, Cat# 7074P2 or Sigma Cat# AP187P) was diluted 1:5000 in blotto. Goat anti-mouse-HRP (Thermo Fisher Cat# 31430) was diluted 1:5000 in blotto.

For total protein detection, Sypro Ruby (Thermo Fisher) was used according to the manufacturer's directions.

4.8.1. Quantitative western blotting

Cell culture

To determine the levels of Pde11 and Raf protein after Pum RNAi, knockdown of Pum protein was performed in Pum-myc/V5-Pde11 or Pum-myc/V5-Raf cells, using 10

µg Pum dsRNA or LacZ dsRNA in a 4-day experiment. Three wells of each condition were used. Cells were plated in 6-well plates at a density of 1 x 10^6 cells/ mL for 4-day knockdowns. Cells were incubated in 10 µg dsRNA targeting Pum for 60 minutes in 1 mL serum-free media. 2 mL complete media was added after dsRNA bathing step and cells were incubated at 25°C for 90 hours.

Western blotting

At harvest, media was removed and cells were washed in 1mL PBS. Cells were resuspended in 150µL RIPA buffer with 2x cOmplete protease inhibitors (Roche). Cells were incubated on ice for 10 minutes then mechanically disrupted with a pestle for 20 seconds each. Lysates were collected after a 10 minute 21,000 x g centrifugation step at 4°C. Protein concentration was determined by Lowry DC with BSA standard (BioRad).

To determine the linear range of detection for V5-tagged proteins, 5 µg-45 µg dilutions were used. For quantitation of Raf and Pde11 protein levels, 5 to 25 µg was used for the LacZ samples and 5 to 20 µg was used for Pum RNAi samples. For each replicate, titrations of total protein were loaded on a 12% SDS-PAGE gel. Gels were electrophoresed at 150 volts until the dye front reached the bottom of the gel, then transferred to Immobilon PVDF membrane overnight at 4°C at 35 volts. Membranes were dried thoroughly, then Sypro Ruby stain was applied following the manufacturer's directions and detected using the auto-exposure setting on the ChemiDoc. After imaging, blots were rocked in 3 changes of blotto for 1 hour to remove residual stain. Mouse anti-V5 antibody (Sigma) was applied and incubated overnight at 4°C. Primary antibody was washed with 3 changes of blotto, and HRP-labeled goat anti-mouse antibody was applied and incubated for 1 hour at room temperature. Secondary antibody was washed with two 5-minute washes of blotto followed by two 5-minute washes with 1x TBS. Blots were incubated briefly in Immobilon substrate (Millipore), quickly rinsed with PBS, and imaged using the optimal auto-exposure chemiluminescence setting on a ChemiDoc Touch.

Data analysis

To determine the linear range of detection, acquired images were analyzed using Bio-Rad Image Lab software. Titrations of 5 µg to 45 µg each of Pde11-V5 and Raf-V5

whole-cell lysates were analyzed via western blots. Three sets of graphs were generated to determine sensitivity range. 1) To measure the sensitivity of Sypro Ruby stain, we used adjusted total band volume for Sypro Ruby stain (described for housekeeping bands below) and graphed it versus total protein loaded. 2) For the V5 antibody, we used adjusted total band volume for V5 antibody and graphed it versus total protein loaded. 3) To ensure that normalizing V5 antibody to Sypro Ruby would fall within the linear range of detection, we also measured normalized V5 antibody band volume (normalized to Sypro Ruby) and graphed it versus total protein loaded. For both Pde11 and Raf detection, 5 to 25 µg total protein lies within the linear range of detection.

To determine protein level fold change, normalized band volume for each loaded protein amount was compared within each blot. For example, the band volume for the 5 µg Pum KD sample was compared to the 5 µg LacZ sample. Fold change was determined as the ratio Pum KD/LacZ. For each individual blot there were a total of 4 measurements (5 µg, 10 µg, 15 µg, 20 µg). For each experiment with 3 biological replicates, this resulted in a total of 12 measurements, and with 3 experimental repeats, a total of 36 measurements were made. The average of these measurements is represented as a log2 scale value, with error bars as standard error of the mean. To generate the standard curve, the adjusted band volume for V5 detection of each µg amount loaded in each experiment was graphed against total protein loaded.

Image Lab software analysis

Multichannel images were created with the chemiluminescent and Sypro Ruby images. Lanes were detected manually. Bands for chemiluminescence were detected automatically if possible. If manually detected, outer regions of bands were chosen to capture the prominent peak and match the profiles of auto-detected bands. The normalization channel was set to Sypro Ruby using housekeeping bands, and large regions of the lanes in the Sypro channel were selected to represent housekeeping bands. If possible, the entire lane was used. This selection method enabled us to avoid regions of saturation at the bottom of the blots or regions with large transfer imperfections. Each lane on the same blot used the same corresponding region for housekeeping band selection. Background subtraction for the Sypro channel was set to a 70 mm rolling disk. Background subtraction for the chemiluminescence channel was

set to the maximum level possible that did not interfere with band detection. Data for all lanes and bands was then exported to Excel, where graphs were generated. For V5-Pde11, the normalized volume for both bands were added together for graphing.

4.8.2. CRISPR generation of Pum-Myc, Pum KO lines, V5-Pde11, V5-Raf

Design of CRISPR guides

For all CRISPR workflows, the NCBI database was used to determine the ideal placement for CRISPR-tagging or gene knockout. If available, the N-terminus of each protein was chosen so that indels were more likely to result in a knockout. Benchling software (2019) was used to identify CRISPR sites with the highest potential efficiency for Cas9 cleavage with the lowest potential off-target effects. Gene-specific primers (listed in Appendix G) were used to amplify the region of interest and TA-cloned into the pGEM T-easy vector to generate a library. Ten clones from each library were mini-prepped using Promega Mini-prep kits and sequenced to verify the targeted region matched the NCBI sequence and that the PAM sites were correct. Single-stranded DNA oligos were ordered to match the sgRNA sequence with additional BspQI overhangs for cloning into the pAc-sgRNA-Cas9 vector (Bassett, Tibbit et al. 2014). Oligos were prepared for cloning by taking 50 pmol of each forward and reverse strand, briefly heating them at 90°C, and then slowly cooling them. The annealed oligos were then phosphorylated using T4 polynucleotide kinase (PNK, NEB) and used in a ligation reaction with a BspQI-digested vector.

The *Drosophila* Pum gene is located on the third chromosome and codes for eight different transcripts. There are 5 annotated protein isoforms, the largest of which is 1533 amino acids and approximately 157.5 kDa (via NCBI + FlyBase). It is unclear if there is isoform-specific expression in different fly tissues. The canonical 1533 aa form, A, is exclusively detected in Dmel2 and Dl1 cells. For this work, all references to amino acid positions or nucleotide bases refer to isoform A or transcript variant A.

For the Pum knockout, the first exon common to all isoforms is exon 9. Primers were designed to amplify 463 bp of exon 9 in WT Dl1 cells, and two Cas9 sites were chosen to target. Two clones with homozygous knockouts of Pum were recovered.

All annotated isoforms of Pum share the same C-terminus, so the C-terminal end of Pum was targeted for the addition of a Myc tag in WT Dl1 cells. Primers were designed to amplify a 831 bp region in exon 13 and two sgRNA sites were chosen and sequence verified. A 200 bp Ultramer oligonucleotide sequence was ordered from IDT containing a single Myc tag sequence immediately preceded by a HRV 3C site. The homology repair (HR) template was designed to have at least 50 bp of overlapping sequence distal and proximal to each cut site. Two verified clones homozygous for the Myc tag insertion were isolated. The first carried an in-frame duplication of the HR insertion resulting in two tandem HRV 3C and Myc sites. The second clone had an appropriate insertion resulting in a functional HRV 3C site and Myc tag. The single tagged Pum line was used in all experiments.

Pum-Myc cell lines were used to create V5-tagged Raf and Pde11. The Pde11 gene has 6 transcript variants which code for 4 unique proteins with different N- and C-termini. This work references variant E, where exon 3 is the first common exon for all protein isoforms. Primers were designed to amplify a 412 bp region of intron 2 and exon 3. A single nucleotide polymorphism was detected in the analyzed sequences for Pum-Myc Dl1 cells where both sgRNAs would bind, and the cloning oligos were changed to match. A 200 bp Ultramer was ordered that would insert a V5 tag internally at the N-terminus of all annotated protein isoforms. For isoforms E, G, and C, this would place the V5 tag beginning at position 22, and position 65 for all others. A single V5-Pde11 clone was isolated. This clone had one allele properly tagged. The other allele contained a single nucleotide deletion which changes the methionine at position 57 to a stop codon. There were no noticeable effects on growth or Pde11 transcript levels with this heterozygous (functionally haploid) line.

The *Drosophila* Raf gene encodes two transcript variants, both of which code for identical proteins. Primers were designed to amplify a 479 bp region of exon 3 of variant E in Pum-Myc Dl1 cells. A 200 bp Ultramer was designed to insert a N-terminal V5 tag at position 2. A single homozygous V5-knock-in clone was isolated and verified by sequencing and western analysis.

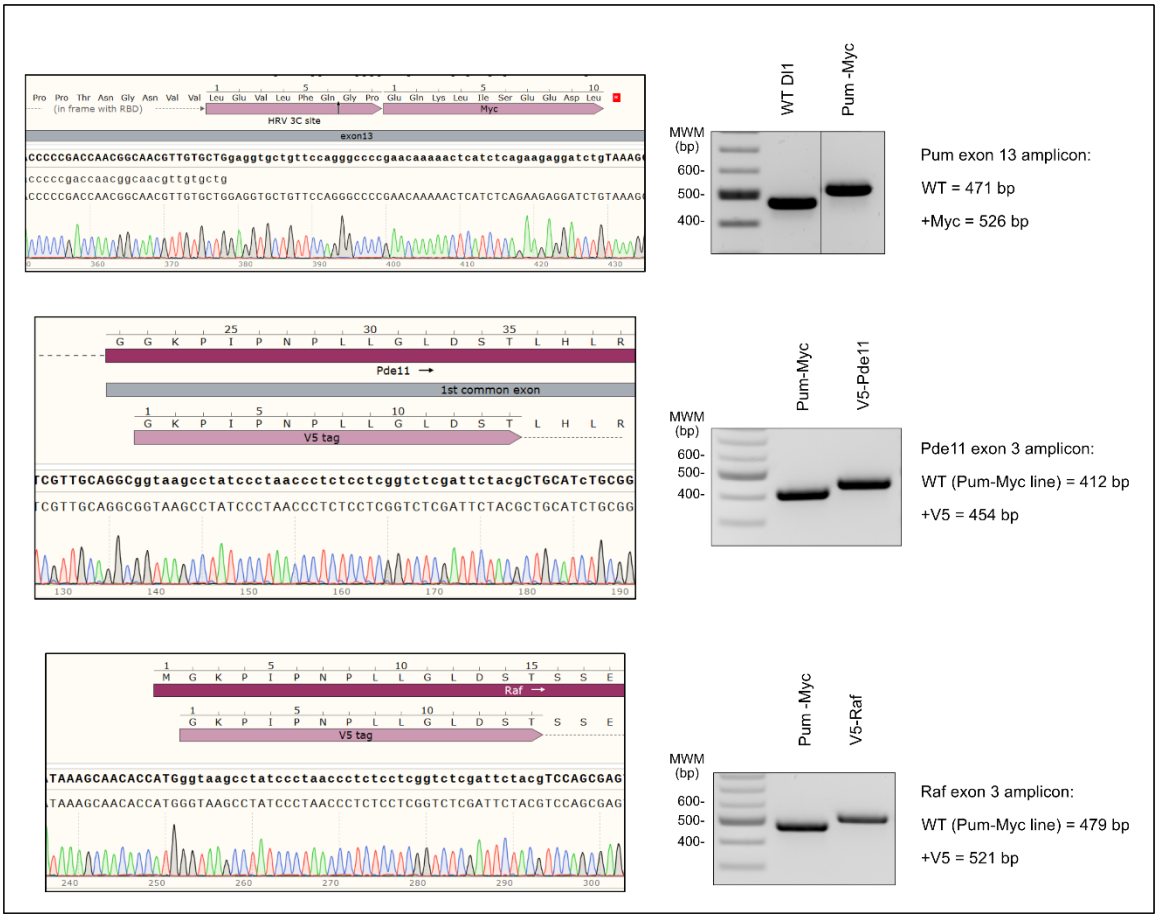


Figure 4.7: Characterization of Pum-Myc, V5-Pde11, and V5-Raf cell lines. Aligned chromatogram of sequenced amplicons using SnapGene v5.0.4. The inserted tags are shown. (Top) C-terminal tagged Pum-Myc with DNA agarose gel of PCR amplicons. (Center) N-terminal V5-tagged Pde11 with DNA agarose gel of PCR amplicons. The single base deletion is further downstream and not pictured. (Bottom) N-terminal V5-tagged Raf with DNA agarose gel of PCR amplicons.

Generation of CRISPR-tagged cell lines

Dl1 cells were plated at a density of 1 x 10^6 per mL in 2 mL complete Schneider's *Drosophila* Medium (SDM) in a six-well dish. Transfection mixes were set up using 4 µL FuGene HD (4:1 ratio)(Promega), 1 µg sgRNA-Cas9 plasmid DNA, 40 pmol homologous recombination (HR) template, and 100 µL serum free SDM. After a 10 minute incubation at room temperature, transfection mixes were applied dropwise to wells. One well remained untransfected to measure efficacy of Puromycin or Blasticidin treatment. Cells were incubated at 25°C for 48 hours, then media was removed and

replaced by media containing 5 µg/mL Puromycin or Blasticidin (Gibco). After 48 to 72 hours, when untransfected cells had died, selective media was removed and replaced with complete SDM. When cells became confluent, they were passaged up to a 10-cm dish and allowed to continue growing to 80% density again. At this stage, cells were collected as a polyclonal population and approximately half were frozen and stored in liquid nitrogen. The remaining cells were diluted to isolate clones in two ways: a limiting dilution was made and cells were plated at a density of 1 to 2 cells per 100µL into 3 to 6 96-well plates, and 10 µL of this dilution was plated into 10-cm plates. Once visible colonies grew in the 10-cm plates, they were carefully pipetted out and transferred to single wells of a 24-well plate. For all wells of clonal isolates, once cells had achieved 80% density, they were passaged up sequentially to larger wells. Once cells had reached confluency in a 12-well plate, a portion were harvested for genomic analysis. Genomic analysis was performed by isolating genomic DNA (gDNA) from the sample using the Wizard gDNA isolation kit (Promega) and using gene-specific primers to amplify the region of intended template insertion. PCR products were analyzed on a 1% agarose gel for correct size, then purified using the Zymo DNA Clean & Concentrator 25 kit, and verified by DNA sequencing. If contamination of wild type cells was observed, cells underwent isolation procedures again until the remaining population appeared to be homozygous for the intended insertion. For Pde11-V5 cells, one allele is properly tagged but the other contains a single nucleotide deletion resulting in a knockout.

Generation of Pum null cell lines

Dl1 cells were plated at a density of 1 x 10^6 per mL in 2 mL complete media in a 6-well dish. Transfection mixes were set up using 8 µL FuGene HD (4:1 ratio), 2 µg sgRNA-Cas9 plasmid DNA, and 1 mL serum free SDM. After a 10 minute incubation at room temperature, transfection mixes were applied dropwise to wells. Cells were incubated at 25°C for 48 hours, then media was removed and replaced by media containing 5 µg/mL Puromycin. Cells were isolated in a similar manner as for CRISPR tagging, however, Puromycin selection was kept on the cells for twenty days to enforce gene knockout. Once clones were isolated and gDNA was sequenced, Synthego ICE analysis was used to determine efficacy of gene knockout.

Characterization of Pum null cell lines

Two unique homozygous knockout lines were constructed: Pum KO1 and Pum KO2. Clone 1 has a 20 bp deletion with an additional 81 bases added, resulting in a net gain of 61 bases. This causes a frame shift after glutamine 725 and truncates the remaining protein sequence at amino acid 733. Clone 2 has a 10 bp deletion in exon 9, causing a frameshift after methionine 726 and truncates the protein at amino acid 764. These lines were created prior to the Pum-Myc line, and therefore cannot be detected on a western blot.

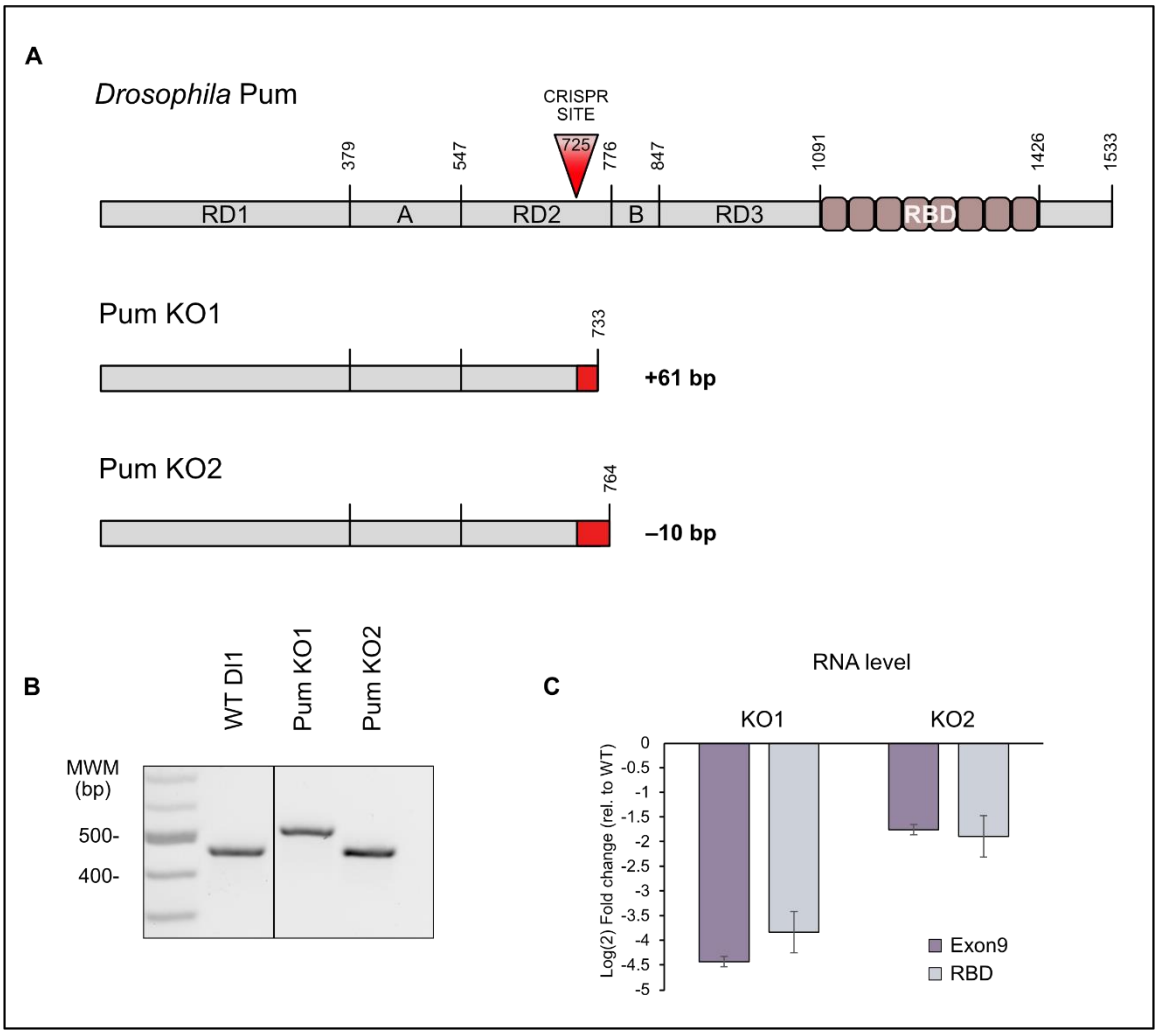


Figure 4.8: Characterization of Pum KO lines. (A) The full-length Pum protein with CRISPR targeting site indicated. The resulting protein products for each Pum KO line are shown, with out-of-frame amino acid sections shown in red. Positions of stop codons are marked, along with the net gain or loss of nucleotides resulting from CRISPR indels. (B) DNA agarose gel showing PCR amplicons from genomic DNA of exon 9. The primers amplify across the CRISPR target region. WT Pum amplicon is 463 bp, Pum KO1 is 524 bp, and Pum KO2 is 453 bp. (C) Pum RNA level in each KO line as measured by RT-qPCR. qPCR primers bind to either exon 9 (left bar) or the RBD (right bar).

Figure 4.9: Chromatogram alignment of Pum KO lines. Exon 9 amplicons from WT Dl1 cells and Pum KO lines were sequenced and aligned to WT Pum using SnapGene v5.0.4. (Top) Pum KO1 has a large insertion and small deletion with a net gain of 61 bases. (Bottom) Pum KO2 has a 10 base deletion.

Chapter 5: Genome-wide analysis of Pum targets

The RNA seq and data analysis in this chapter were performed in collaboration with the laboratory of Dr. Eric Wagner at the University of Texas Medical Branch including Nathan Elrod (PAC-Seq and Click-Seq data analysis), Madeline K. Jensen and Ping Ji (library generation).

5.1. Introduction

The repertoire of mRNAs that are functionally regulated by *Drosophila* Pum has not been analyzed on a global scale. Previous work focused on select mRNAs linked to phenotypes in the germline, embryo, and nervous system (Gerber, Luschnig et al. 2006, Laver, Li et al. 2015). As described in the preceding chapter, analysis of the mRNAs associated with Pum indicated the potential for a much broader regulatory role, with many hundreds of transcripts bound to Pum. In model organisms (yeast, *C. elegans*, mouse) and human cells, PUMs regulate hundreds of mRNAs (Wickens, Bernstein et al. 2002). Thus, we set out to analyze the effect of Pum and its CNOT corepressors on the *Drosophila* transcriptome.

In this chapter, we used two strategies to identify the mRNAs regulated by Pum in the DL1 cell line. DL1 cells were chosen because they are amenable to transfection, support efficient RNAi, and are facile for CRISPR-Cas9 genome engineering and clonal isolation. In the first approach, we performed RNAi to transiently deplete Pum protein. In the second approach, we used CRISPR-Cas9 editing to create several clonal DL1 lines with indels that knock out Pum.

While RNAi does not provide a null background, it is rapid and transient and therefore may be less prone to adaptive changes. CRISPR-Cas9 KO offers a complete loss of function but may be subject to adaptive changes in the cells during the length of selection and clonal isolation. Here we use a combination of both approaches to rigorously analyze Pum regulation.

We recently showed that the CNOT complex is necessary for Pum repression of PRE-containing reporter genes. RNAi of Pop2 or Not1 reduced translational repression and degradation of PRE-containing reporter mRNAs (Arvola, Chang et al. 2020). In

Chapter 4 we demonstrated that depletion of Not1 increased the levels of endogenous Pum targets. In this study, we extend this approach to analyze the impact of Not1 and Pop2 depletion on the transcriptome. We then compare the mRNAs regulated by Pum and CNOT as a means to evaluate the impact of the Pum-CNOT axis on the transcriptome in flies.

5.2. Pum-CNOT mediated regulation of the transcriptome using PAC-Seq

To identify mRNAs that are differentially regulated in response to depletion of Pum, we used the recently developed form of RNA sequencing, PAC-Seq (Poly-A-Click-Seq)(Routh, Ji et al. 2017). Library preparation for PAC-Seq consists of annealing an oligo dT primer containing a portion of the Illumina p7 adapter to RNA transcripts, then performing reverse transcription in the presence of azido-modified A, G, and C nucleotides which cause the reaction to stall upstream of the poly(A) tail. Click-chemistry is then used to ligate the Illumina p5 adapter to the 3' end of the first strand. PCR amplification is performed using a universal primer and an indexing primer, then the library is size selected on an agarose gel to retain amplicons 200 to 400 bp in length. PAC-seq has several advantages over traditional RNA sequencing methods: there is no need to enrich for poly(A) tails or deplete rRNA, RT termination and size selection eliminate the need for RNA fragmentation, and click chemistry avoids the inefficient step of linker ligation. In addition, PAC-Seq requires only 2 µg of total RNA.

To prepare samples for PAC-Seq, we used our knockdown approach to deplete Pum, Not1, and Pop2 using RNAi in Pum-Myc Dl1 cells. An optimized protocol was used that included treatment of cells with dsRNAs to ensure efficient depletion of the target proteins. A non-targeting control dsRNA corresponding to *E.coli* LacZ was used for normalization. Three biological replicates were analyzed for each condition. Depletion of Pop2 and Not1 proteins by these dsRNAs was previously reported (Arvola, Chang et al. 2020) and was confirmed in these samples by RT-qPCR (Figure 5.1 B). Depletion of endogenous myc-tagged Pum protein was confirmed by western blotting and RT-qPCR (Figure 5.1 A, B).

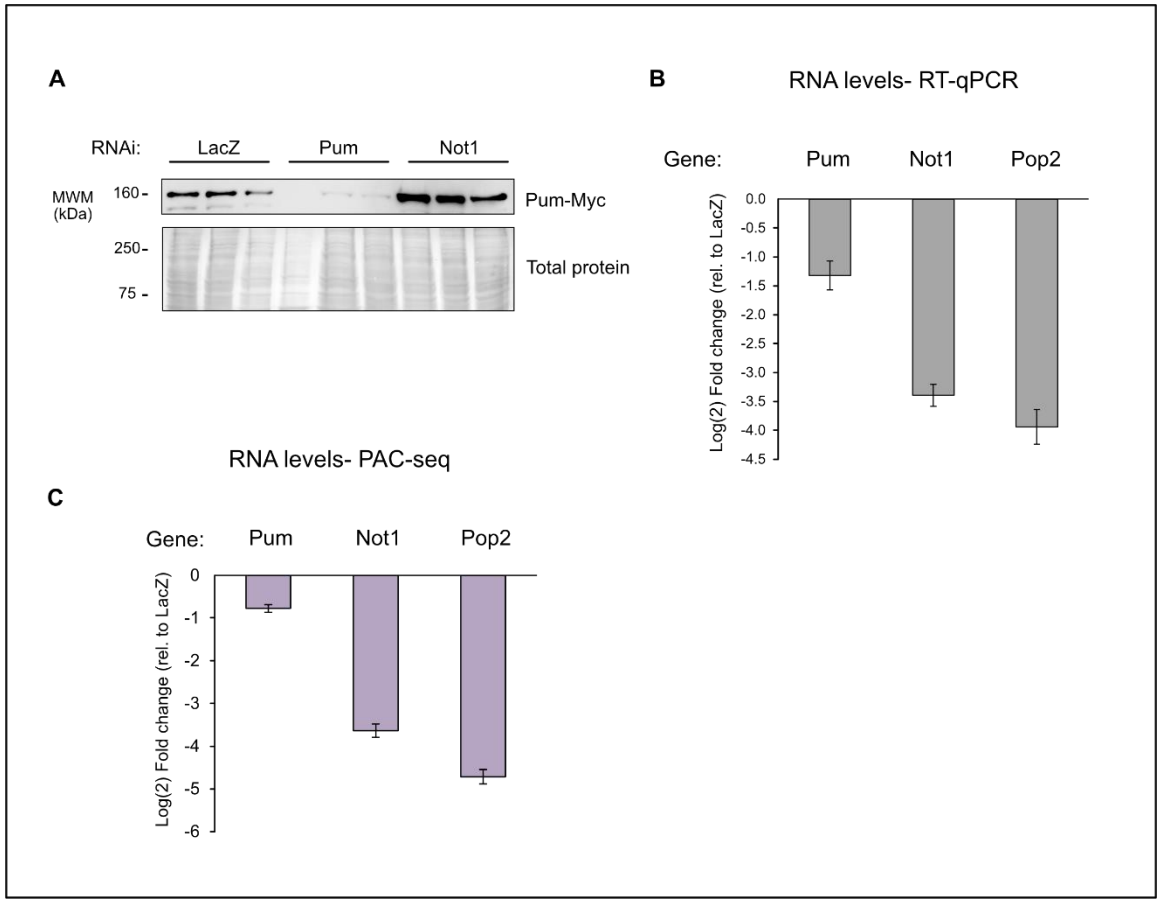


Figure 5.1: RNA levels of Pum, Not1 and Pop2 used in sequencing experiments. (A) Western blot showing depletion of Pum-Myc in Dl1 cells. Total protein is shown as a loading control. Not1 and Pop2 depletion cannot be measured by western. (B) Mean RNA levels of Pum, Not1, and Pop2 as measured by RT-qPCR (n = 9). Log(2) fold change is shown relative to LacZ. Error bars represent standard error of the mean (SEM). (C) Mean RNA levels of Pum, Not1, and Pop2 relative to LacZ, as measured by PAC-seq. Error is SEM.

A total of 9,284 genes were measured with a range of 0.02 to 165,116 reads per kilobase million (RPKM). We first trimmed this data set to exclude transcripts from the mitochondrial genome, pseudogenes, and genes that had withdrawn gene status as of FlyBase release 6.38 (February 2021). Other transcripts which could not be mapped specifically to a protein-coding gene, ncRNA, miRNA, or annotated miscellaneous RNA were also eliminated. After trimming, 9,262 unique genes remained.

The PAC-Seq data confirm depletion of Pum, Not1, and Pop2 (Figure 5.1 C). Not1 and Pop2 showed high levels of depletion, with a 12.4-fold decrease for Not1 and

26.2-fold decrease for Pop2. The level of depletion for Pum (1.7-fold decrease) was less robust. It is worth noting, however, that this level of Pum depletion resulted in substantial stabilization of PRE-containing reporter mRNAs (Arvola, Chang et al. 2020) and alleviation of Pum repression in Dmel2 cells (Weidmann and Goldstrohm 2012, Arvola, Chang et al. 2020).

5.2.1. Genes affected by Pum depletion

To explore which genes were affected by depletion of Pum, we established a biological significance threshold of +/- 1.3-fold change. This change corresponds to the level of repression seen in our reporter system in human cells when a single PRE is used in a minimal 3'-UTR (Bohn, Van Etten et al. 2018). For statistical significance calling we used a p-adjusted value of 0.05.

Of the 9,262 genes measured in our PAC-Seq experiment, only 79 were significantly upregulated in the Pum RNAi condition (Figure 5.2 A). We performed gene ontology (GO) analysis on this set of genes. The most significantly enriched terms include protein localization, establishment or maintenance of cell polarity, motor neuron axon guidance, cell to cell adhesion, and the Notch signaling pathway. Most of these genes also contain one or more perfect PREs (Figure 5.2 B). Approximately 28% of all the genes measured contain PREs (2,629 out of 9,262), yet 75% of the upregulated genes contain PREs (51 out of 79), indicating that Pum is causing specific regulation of the PRE-containing genes (odds ratio 4.67, 95% Credible interval = 2.94, 7.42).

Knockdown of Pum also caused a decrease in the RNA level of 51 genes, a minority of which contain PREs (Figure 5.2 B). We performed GO analysis on these genes and the significantly enriched categories are all involved in stress responses such as heat shock, protein unfolding, and hypoxia.

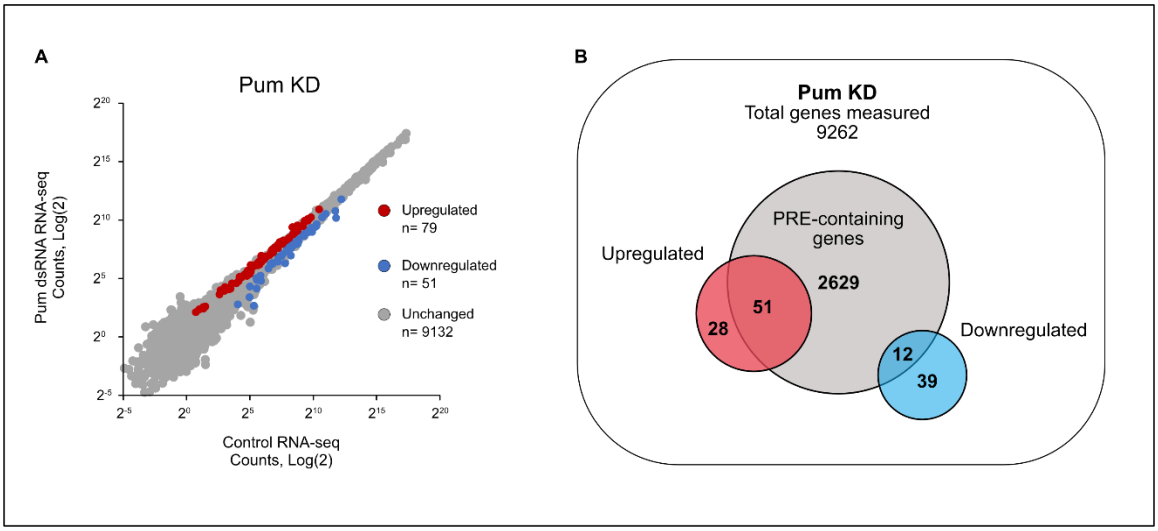


Figure 5.2: Genes affected by Pum depletion. (A) Plot showing PAC-seq reads of Pum RNAi condition versus Control LacZ. Significantly upregulated genes are shown as red points, while downregulated genes are shown in blue. (B) Venn diagram representing the total number of genes measured, PRE-containing genes, and Pum regulated genes.

5.2.2. Genes affected by CNOT complex depletion

Next, we examined the genes affected by Not1 and Pop2 depletion. Since the CNOT complex is the major cytoplasmic deadenylase complex in *Drosophila*, we expected a large overall impact on the transcriptome. We used the same biological and statistical significance threshold as for Pum targets in our analysis. Over 2000 genes are upregulated when either Not1 or Pop2 are knocked down, with 1,790 of these genes upregulated in both conditions (Figure 5.3). Additionally, over 2000 genes are downregulated, with 1789 shared. Of the genes that were not similarly regulated between Not1 and Pop2, most were affected alike but missed our cutoffs for fold change (> 1.3-fold) or statistical significance. In fact, only two genes were differentially regulated, *CDC45L* and *CG6231*. *CDC45L*, a cell cycle gene, was decreased in the Not1 knockdown and increased in the Pop2 condition. *CG6231* encodes a protein that is predicted to be involved in transmembrane transport. It was increased in the Not1 knockdown but significantly decreased for Pop2.

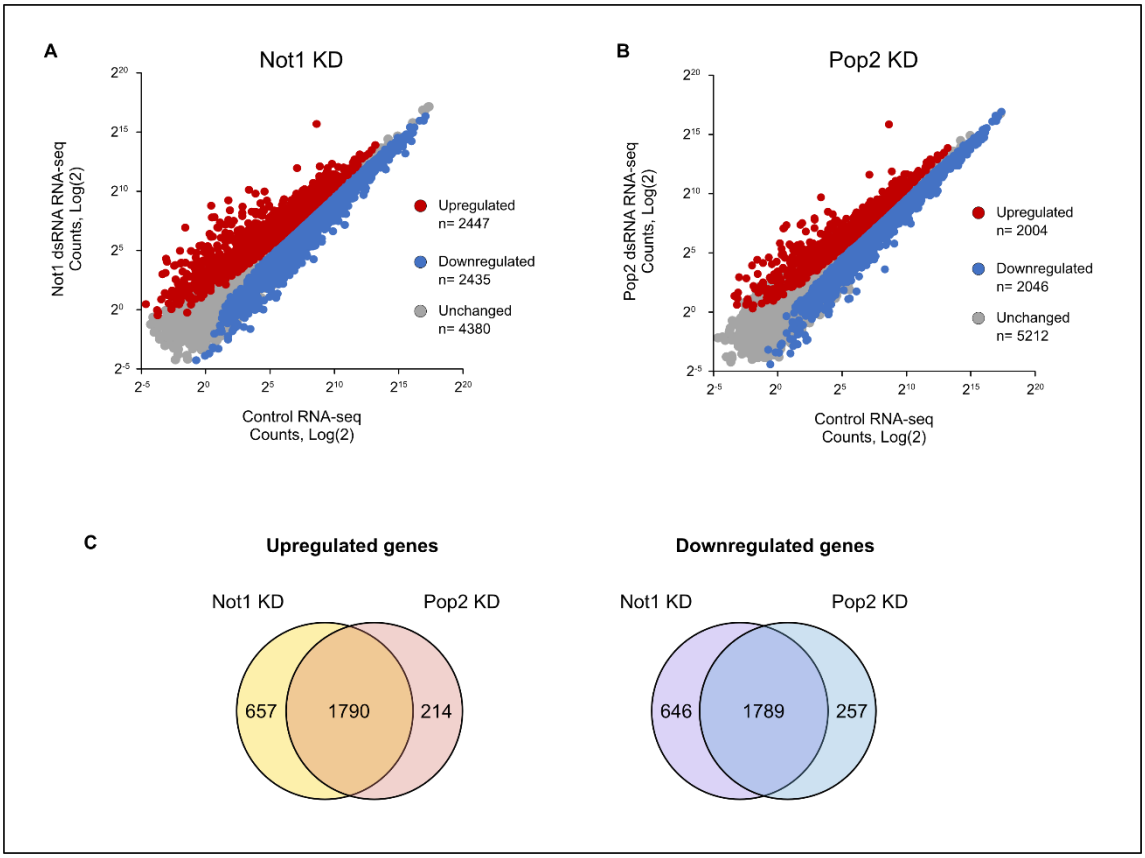


Figure 5.3: Genes affected by knockdown of Not1 and Pop2. (A) Plot of total genes measured in Not1 RNAi condition as log(2) counts of control condition versus Not1. Significantly upregulated genes are shown as red points, while downregulated genes are shown in blue. (B) Plot of total genes measured in Pop2 RNAi condition as described for Not1. (C) Venn diagrams illustrating overlap of upregulated genes between Not1 KD and Pop2 KD.

We performed GO analysis on the 1790 shared upregulated targets for Not1 and Pop2. As expected, the enriched terms cover a wide array of biological processes and molecular functions. The most significant terms included protein binding, intracellular protein transport, Rab GTPase binding, vesicle-mediated transport, and border follicle cell migration. The top Kyoto encyclopedia of genes and genomes (KEGG) pathway results were endocytosis, DNA replication, and SNARE interactions in vesicular transport. Consistent with the CNOT complex's role in RNA regulation, GO analysis on the downregulated targets showed highly enriched categories in processes such as cytoplasmic translation, structural constituent of ribosome, and RNA binding. KEGG pathways enriched included the ribosome pathway and oxidative phosphorylation.

Because Pum requires the CNOT complex to repress reporter genes, we predicted that the genes upregulated in the Pum knockdown condition should overlap with the Not1 and Pop2 knockdowns. We compared all three datasets for both upregulated and downregulated genes. Just under half of the genes in the Pum downregulated gene sets overlapped for the 3 RNAi conditions. We interpret this to represent shared secondary effects that may result from loss of Pum and CNOT function. There is a pronounced overlap between the genes in each condition, with 40 genes upregulated by depletion of Pum, Not1, and Pop2 (Figure 5.4).

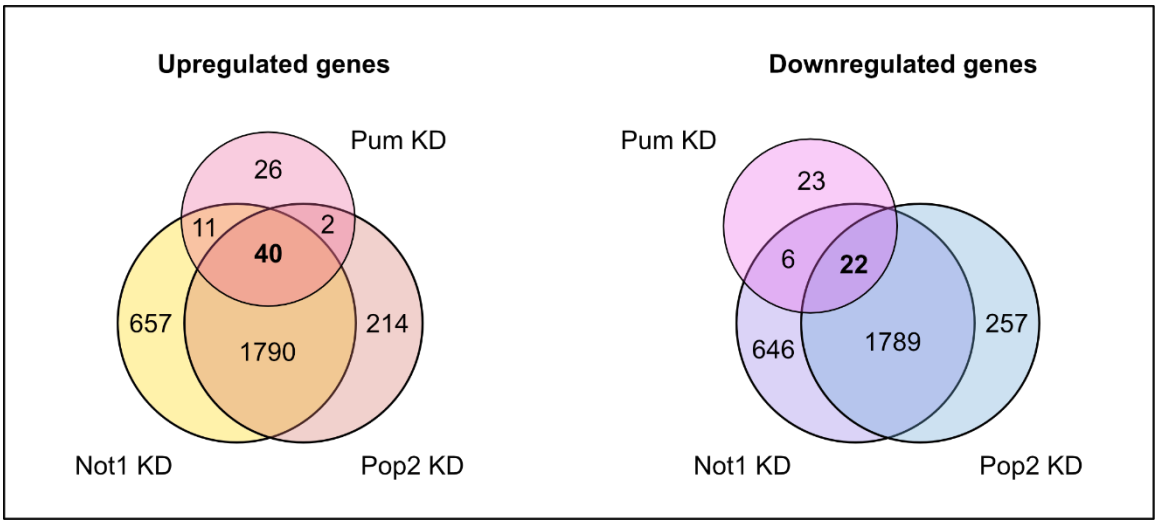


Figure 5.4: Overlap of upregulated and downregulated genes between Pum, Not1, and Pop2 in the PAC-Seq experiment.

Of the 40 shared upregulated genes, 25 contain at least one perfect PRE, indicating they could be direct Pum target mRNAs. Further support is provided by RIP-Chip data, which provide experimental evidence for Pum-binding to these mRNAs (Gerber, Luschnig et al. 2006, Laver, Li et al. 2015). These genes are listed in Table 5.1. Together, this data supports the importance of CNOT for Pum degradation of PRE-containing mRNAs.

Gene name	FlyBase ID	# PREs	Pum KD log(2) FC	Binding data	Human homologs
wb	FBgn0261563	4	1.39		LAMA(1,2,3,5)
bib	FBgn0000180	2	1.31	(2)	AQP(1-6,8,9), MIP
beat-IV	FBgn0039089	1	1.24		
fz2	FBgn0016797	1	1.18		FZD(5,8), FRZB, SFRP4
CG2150	FBgn0003065	1	1.05		
Msr-110	FBgn0015766	1	1.05		
fend	FBgn0030090	5	0.79	(2)	
Tet	FBgn0263392	1	0.76		TET(1,2,3)
CG34347	FBgn0085376	3	0.66	(1)	EPB41L4A
Ppox	FBgn0020018	1	0.66	(1), (2)	PPOX
fax	FBgn0014163	1	0.63		FAXC
CG11110	FBgn0034535	1	0.62		IMMP1L, IMMP2L
CG2200	FBgn0030447	2	0.55	(1), (2)	
pyd	FBgn0262614	2	0.54	(1)	TJP(1,2,3)
mlt	FBgn0265512	1	0.54		TBCEL
bbg	FBgn0087007	2	0.54		IL16
CG5056	FBgn0032231	2	0.54	(1), (2)	OSER1
numb	FBgn0002973	1	0.51	(2)	NUMB, NUMBL
gukh	FBgn0026239	1	0.47		NHS
COQ7	FBgn0029502	1	0.46	(1), (2)	COQ7
Snx1	FBgn0031534	1	0.45	(1), (2)	SNX(1,2,30)
Atf3	FBgn0028550	3	0.45		ATF3, FOS, FOSB, FOSL(1,2), BATF(1-3), JDP2
Eip93F	FBgn0264490	5	0.45		LCOR, LCORL
CG45263	FBgn0266801	3	0.42		
RnrS	FBgn0011704	1	0.40	(1), (2)	RRM2, RRM2B

Table 5.1: PRE-containing genes affected by Pum KD and overlapping with Not1 and Pop2.

5.3. Pum-CNOT mediated regulation of the transcriptome using Click-Seq and Pum knockout cell lines

The inefficient knockdown of Pum in the preceding experiment was a potential limitation to identification of the full scope of Pum's effects on endogenous genes, including the number of affected genes and the magnitude of the effect. Therefore, we knocked out Pum in DL1 cells using CRISPR-Cas9 to introduce indels into exon 9 that inactivate the gene. We included three replicates each of two independent Pum null cell lines, with WT DL1 cells as a control. The effects of these Pum KOs on gene expression was then compared to Pum RNAi. We further optimized the RNAi conditions to improve depletion of Pum, which we confirmed by RT-qPCR and western blotting. A full characterization of the Pum knockout (KO) lines is provided in Section 4.8.2.

We then performed transcriptome analysis of Pum knockout and knockdown, as well as Not1 and Pop2 knockdowns. To assess reproducibility, 3 biological replicates were performed for each condition. In this experiment, we utilized Click-Seq (Routh, Head et al. 2015). This library preparation method is the parent of PAC-Seq and bears one major difference. Following poly(A) selection to enrich for mRNA, instead of using template reverse primers with oligo dT to target the poly(A) tail of transcripts, Click-Seq uses random primers that can anneal anywhere in the transcript. As our focus was to identify transcripts regulated by factors that act on poly(A) tails, potential library bias toward polyadenylated transcripts was not a concern. In addition, the Click-Seq method had the potential to improve our sequencing depth of coverage.

For the RNAi samples, Click-Seq measured 8,549 total genes with FKPM values from 0.013 to 63969. For the Pum knockout samples, 9,848 genes were measured at FPKM values from 0.017 to 55203. The RNA levels in the Pum KO samples were decreased 2.9-fold and 1.4-fold for KO1 and KO2, respectively (Figure 5.5 C). For the RNAi samples, Pum showed a 3.2-fold decrease, an improvement over the PAC-Seq experiment. We also confirmed Pum-Myc depletion via western blot (Figure 5.5 A). Not1 KD resulted in a 9.6-fold decrease in Not1 RNA levels, while the Pop2 knockdown showed a 20-fold decrease of Pop2 mRNA (Figure 5.5 C).

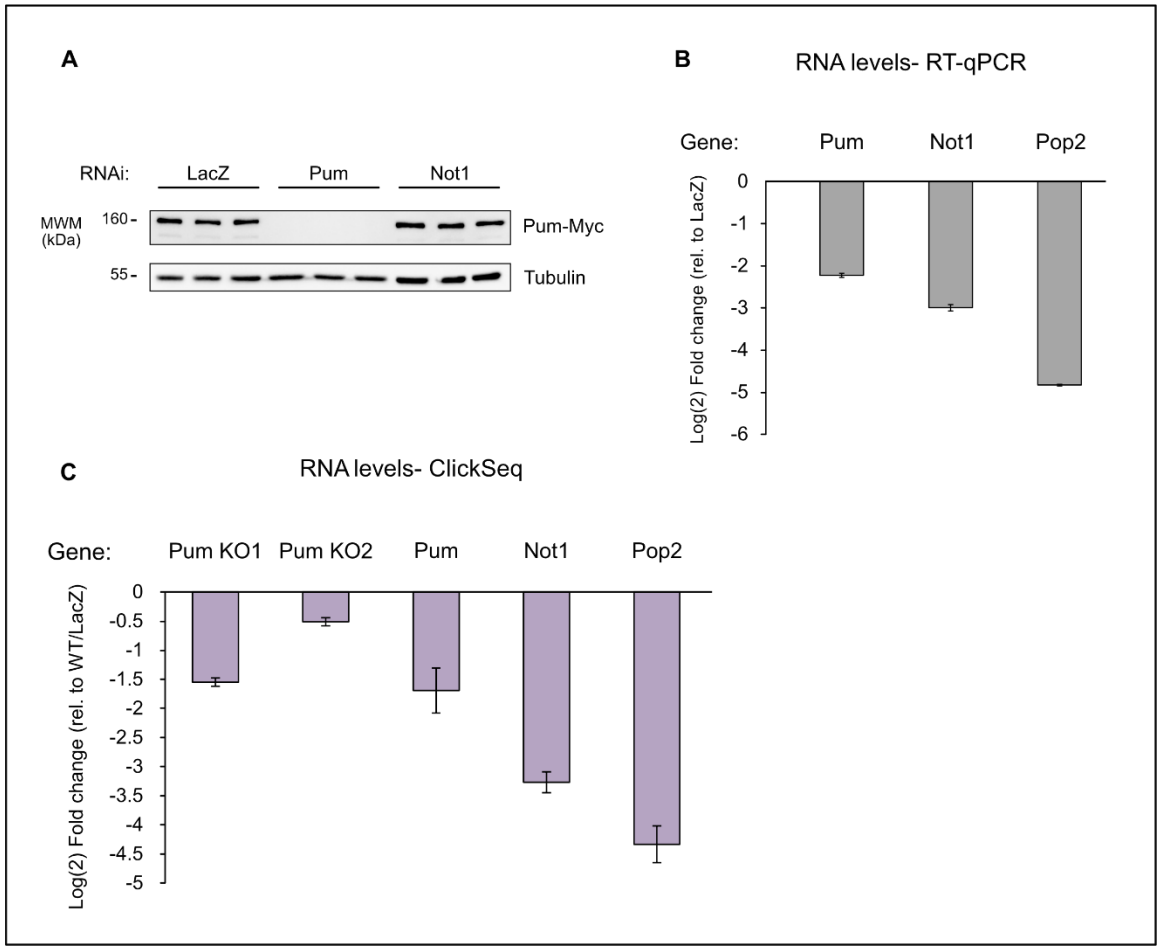


Figure 5.5: RNA levels of Pum, Not1 and Pop2 used in sequencing experiments. (A) Western blot showing depletion of Pum-Myc in DI1 cells. Tubulin is shown as a loading control. Not1 and Pop2 depletion cannot be measured by western. (B) Mean RNA levels of Pum, Not1, and Pop2 as measured by RT-qPCR (n = 9). Log(2) fold change is shown relative to LacZ. Error bars represent standard error of the mean (SEM). (C) Mean RNA levels of Pum, Not1, and Pop2 relative to LacZ, as measured by Click-Seq. RNA levels measured in Pum KO lines are included. Error is SEM.

5.3.1. Genes affected by Pum depletion and knockout

To score differentially expressed genes, we used the same thresholds for biological significance (1.3-fold) and statistical significance (p-adj < 0.05). In the Pum RNAi condition, 44 genes are significantly upregulated for our Click-Seq experiment (Figure 5.6), the majority of which contain PREs. Fourteen genes overlap between the PAC-Seq data and the Click-Seq data for Pum RNAi. Thirteen of these overlapping genes contain PREs. We performed GO analysis on both up- and down-regulated genes

for Pum KD. Notch signaling was again one of the most significantly enriched GO terms in the Pum upregulated samples. Only one term was enriched for the downregulated genes: pyruvate metabolism, represented by two genes, neither of which contain PREs.

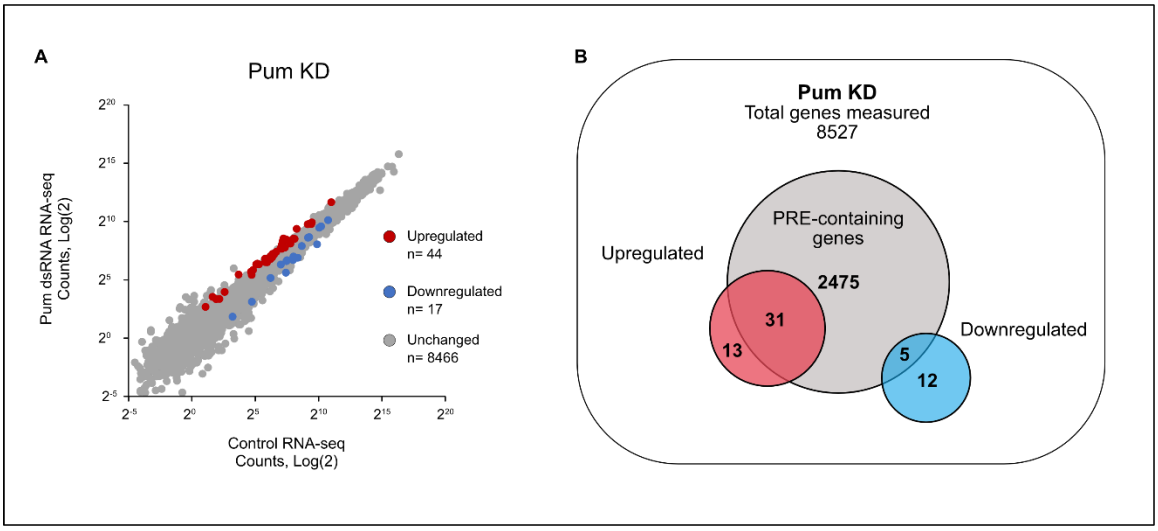


Figure 5.6: Genes affected by Pum depletion in Click-Seq experiment. (A) Plot showing Click-Seq reads of Pum RNAi condition versus Control LacZ. Significantly upregulated genes are shown as red points, while downregulated genes are shown in blue. (B) Venn diagram representing the total number of genes measured, PRE-containing genes, and Pum-regulated genes.

For the Pum KO lines, over 3,000 genes were differentially expressed. There was a strong overlap between Pum KO1 and Pum KO2 for both upregulated and down regulated genes (Figure 5.7). Due to the large number of genes regulated in these samples, we limited our analysis to these overlapping genes.

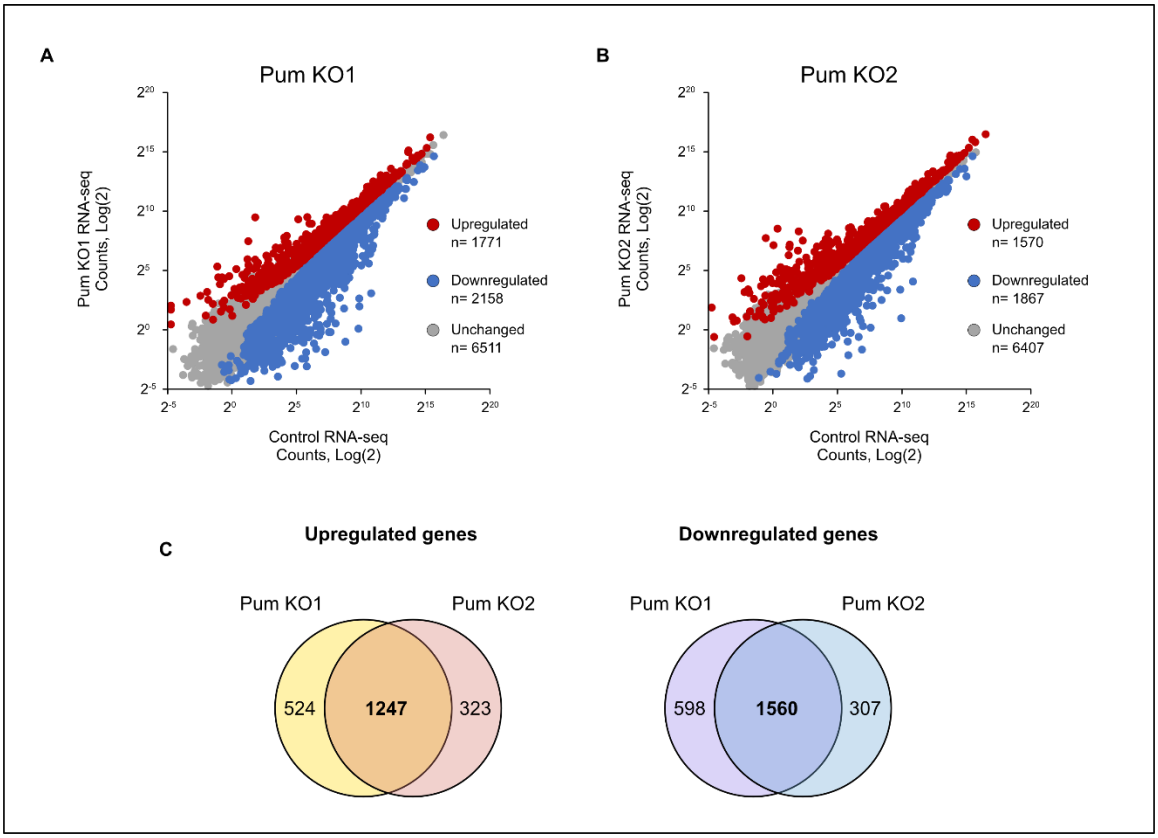


Figure 5.7: Differential expression of genes in Pum KO cell lines. (A) Plot of total genes measured in Pum KO1 lines as log(2) counts of WT versus Pum KO. Significantly upregulated genes are shown as red points, while downregulated genes are shown in blue. (B) Plot of total genes measured in Pum KO2 as described for Pum KO1. (C) Venn diagrams illustrating overlap of upregulated genes between both Pum KO lines.

Of the 1247 upregulated genes, 401 of them contain PREs. We performed GO analysis on both sets (all upregulated and upregulated PRE-containing). Neurogenesis, RNA binding, and mRNA splicing via the spliceosome were some of the highest enriched terms for all upregulated genes (Figure 5.8). Neurogenesis was also one of the top enriched terms for PRE containing genes, along with transcription, axon guidance, protein binding, and the Notch signaling pathway (not pictured).

We then looked at the genes that were downregulated in the Pum KO lines. 465 of the downregulated genes contain one or more perfect PREs. GO analysis on all downregulated RNAs showed enrichment for many terms such as endocytosis, actin binding, neuromuscular synaptic transmission, and oxidoreductase activity (Figure 5.8). When we narrowed our focus to PRE-containing genes only, a pattern of neural related

pathways was prevalent. More than half of the top ten enriched terms pertain to neurons or synaptic transmissions (not pictured).

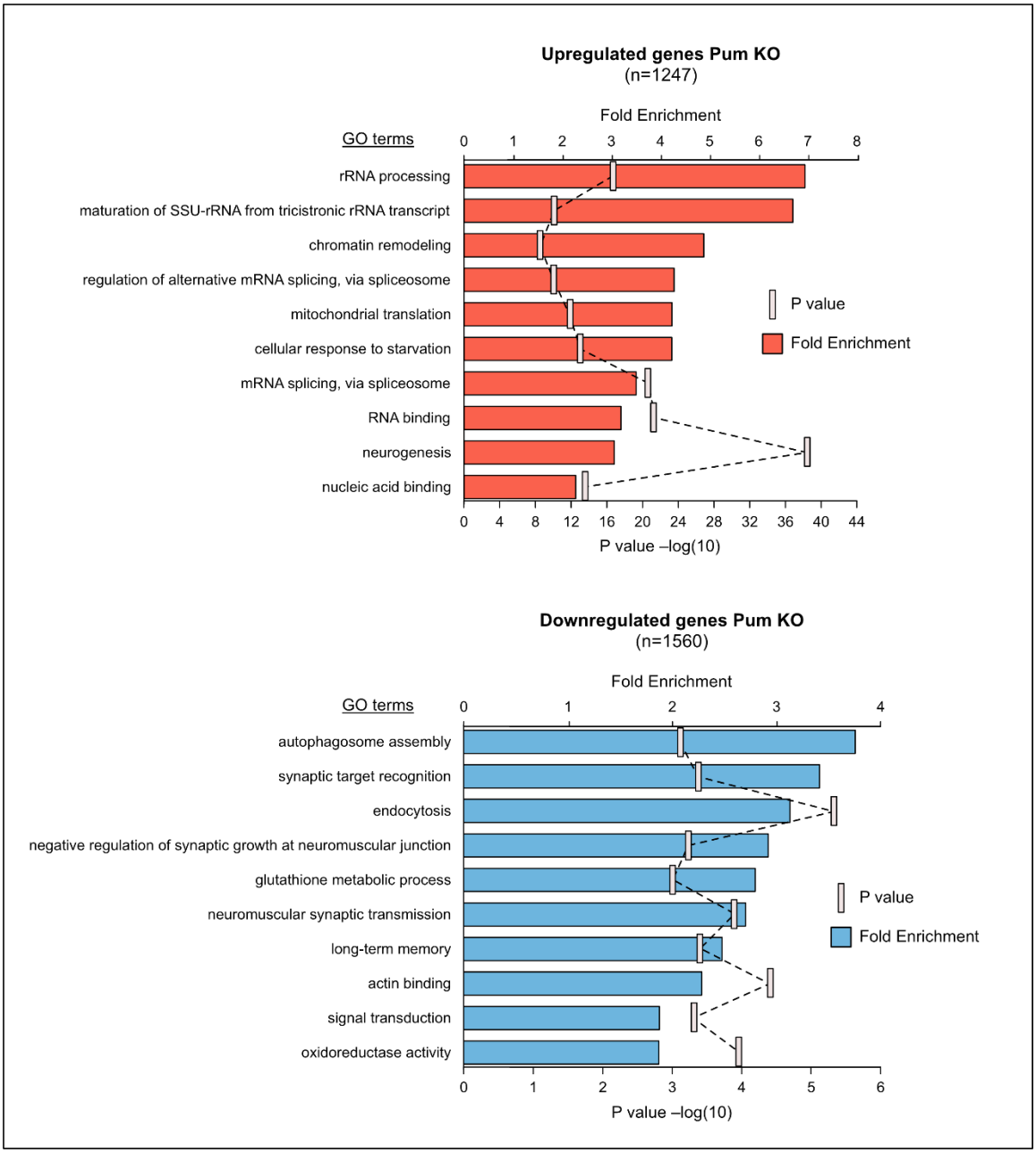


Figure 5.8: Gene ontology analysis of differentially expressed genes in the Pum KO lines. Fold enrichment is shown as full bars corresponding to the upper x axis. P value (as inverse log(10)) is shown as small bars and corresponds to the lower x axis. (Top) GO analysis on significantly upregulated genes. (Bottom) GO analysis on significantly downregulated genes. Data used for the graphs is reported in Appendix F.

5.3.2. Genes affected by CNOT depletion

Similar to our PAC-Seq results, Click-Seq identified thousands of genes that were differentially expressed in response to Not1 and Pop2 depletion, with a clear overlap of genes between both knockdown conditions (Figure 5.9). The total number of genes with changed RNA levels was lower than for our PAC-Seq experiment, which could be an effect of lower Not1 and Pop2 depletion levels. However, the overall ratios of regulated genes and overlap are similar. GO analysis revealed many of the same significantly enriched categories as well.

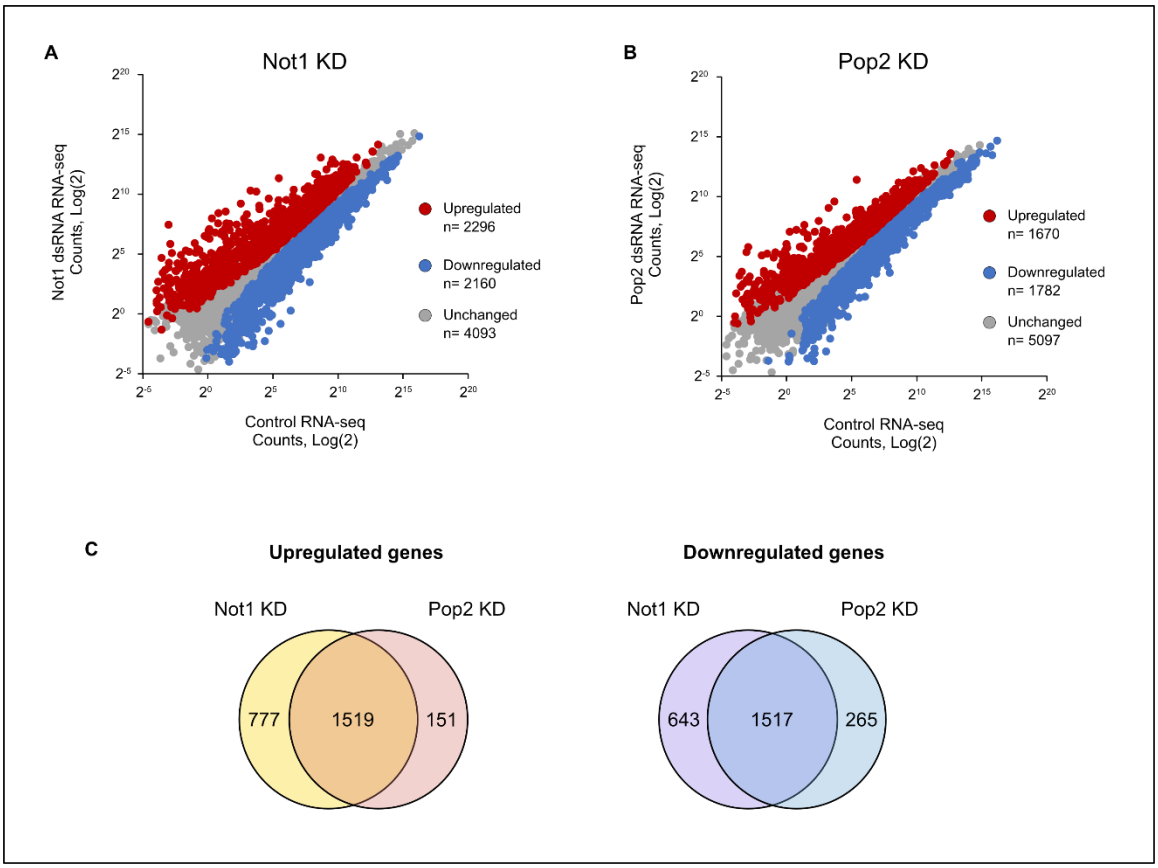


Figure 5.9: Genes affected by knockdown of Not1 and Pop2. (A) Plot of total genes measured in Not1 RNAi condition as log(2) counts of control condition versus Not1. Significantly upregulated genes are shown as red points, while downregulated genes are shown in blue. (B) Plot of total genes measured in Pop2 RNAi condition as described for Not1. (C) Venn diagrams illustrating overlap of upregulated genes between Not1 KD and Pop2 KD.

We compared the genes between both PAC-Seq and Click-Seq for Not1 and Pop2. Overall, 2,467 unique genes are represented in the upregulated category, while 1,932 genes are downregulated. 1037 genes are increased in abundance for both Not1 and Pop2 knockdowns in both datasets, and 1,070 genes are decreased. We performed gene ontology analysis on these genes, as these genes represent our highest confidence CNOT complex regulated transcripts. The effects of CNOT depletion were evident in many different biological processes (Figure 5.10). The GO analysis for downregulated genes stresses the importance of the CNOT complex as a central regulator of translational repression.

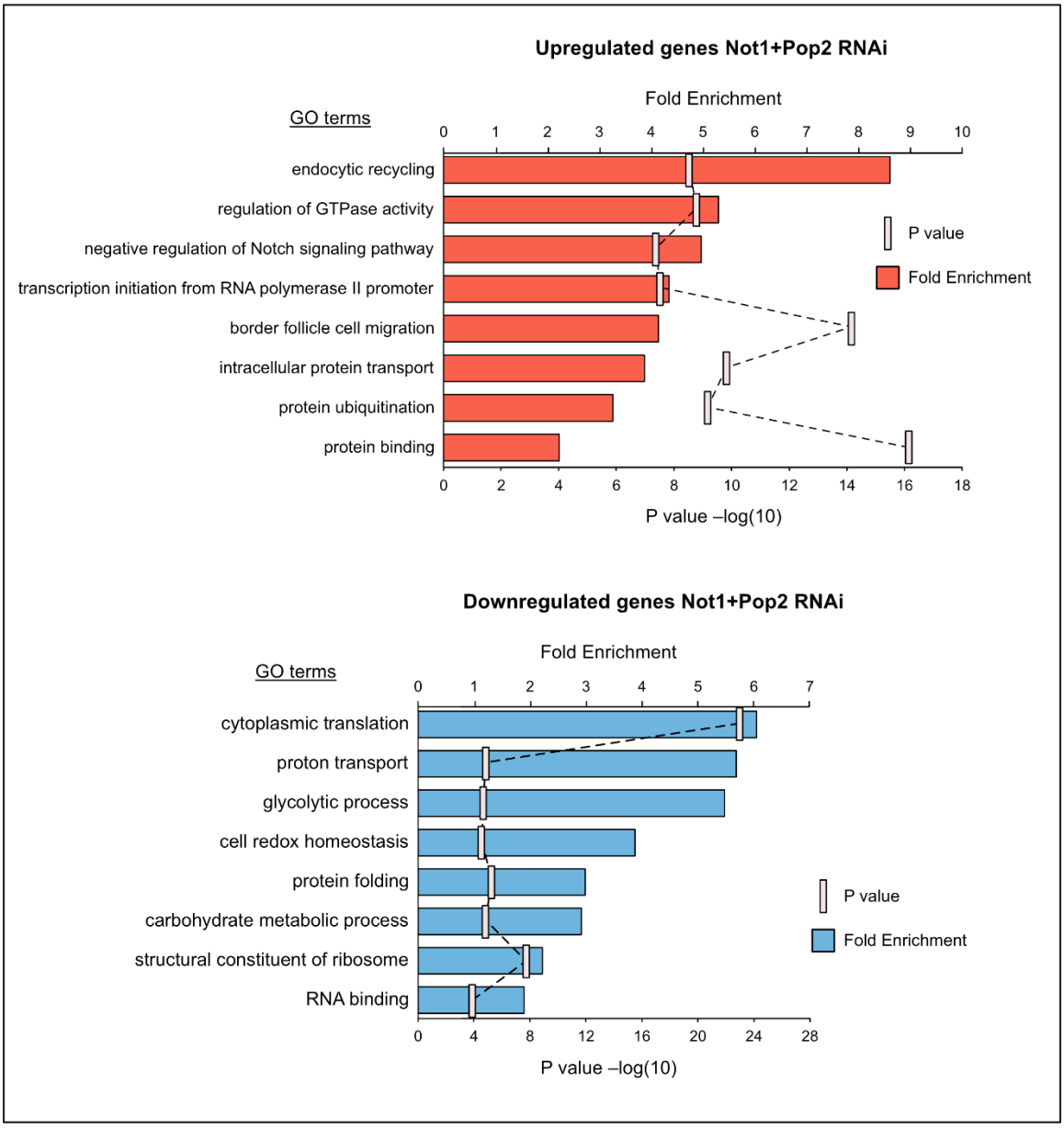


Figure 5.10: Gene ontology analysis for high confidence genes in Not1 and Pop2 RNAi data. Fold enrichment is shown as full bars corresponding to the upper x axis. P value (as inverse log(10)) is shown as small bars and corresponds to the lower x axis. (Top) GO analysis on significantly upregulated genes. (Bottom) GO analysis on significantly downregulated genes. Data used for the graphs is reported in Appendix F.

We then compared the Not1- and Pop2-regulated genes to our Pum-regulated genes in the KO lines. As before, there was some overlap of the genes in each condition (Figure 5.11). Within the overlapping upregulated genes set, 55 of the 121 genes contained PREs, and there were 42 PRE-containing genes in the overlapping downregulated condition. Four of the upregulated genes with PREs were also present in the PAC-Seq upregulated gene list for all conditions: *bib*, *bbg*, *numb*, and *tet*.

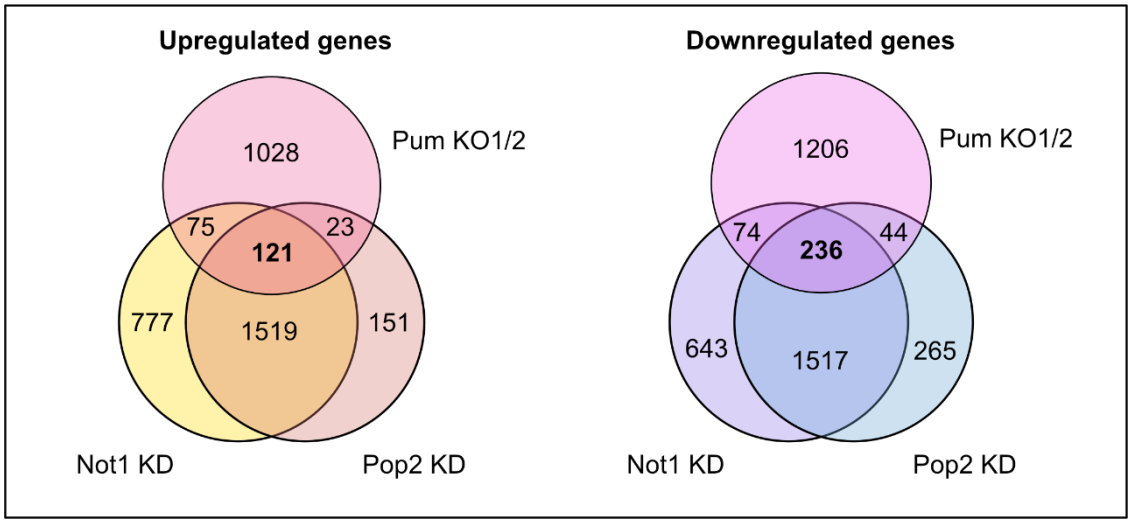


Figure 5.11: Overlap of significantly regulated genes for Pum KO lines, Not1, and Pop2 RNAi.

Another way we evaluated Pum:CNOT targets was by first identifying the PRE-containing genes that were affected by Pum depletion or knockout. There were 416 genes in our Click-Seq data upregulated in either Pum condition. Of these genes, 108 were also upregulated when either Not1 or Pop2 were depleted.

5.4. Analysis of all RNA-seq datasets

Having compiled our data for PAC-Seq and Click-Seq, we next examined the crossover between all the experimental conditions. Since we were focused primarily on identifying Pum targets, we first asked what genes were upregulated in all the Pum KD/KO samples. Across all three experiments (PAC-Seq RNAi, Click-Seq RNAi, Click-

Seq KO), we measured 9,850 unique *Drosophila* genes in total. Of those, 2,144 were significantly upregulated when Pum protein levels were depleted, 2,094 of which were from the KO lines data. These genes were narrowed down further to include only genes which contained one or more PREs and could thus be considered direct Pum targets. This resulted in 696 genes.

The best candidates for direct Pum regulation would presumably not only possess PREs, but would be upregulated in more than one dataset. Of the 696 PRE-containing genes, 417 genes fit this criteria. We performed GO analysis on this set of genes (Figure 5.12). Hemopoiesis and gene silencing were among the most significantly enriched terms, along with chromatin modification, axon guidance and neurogenesis.

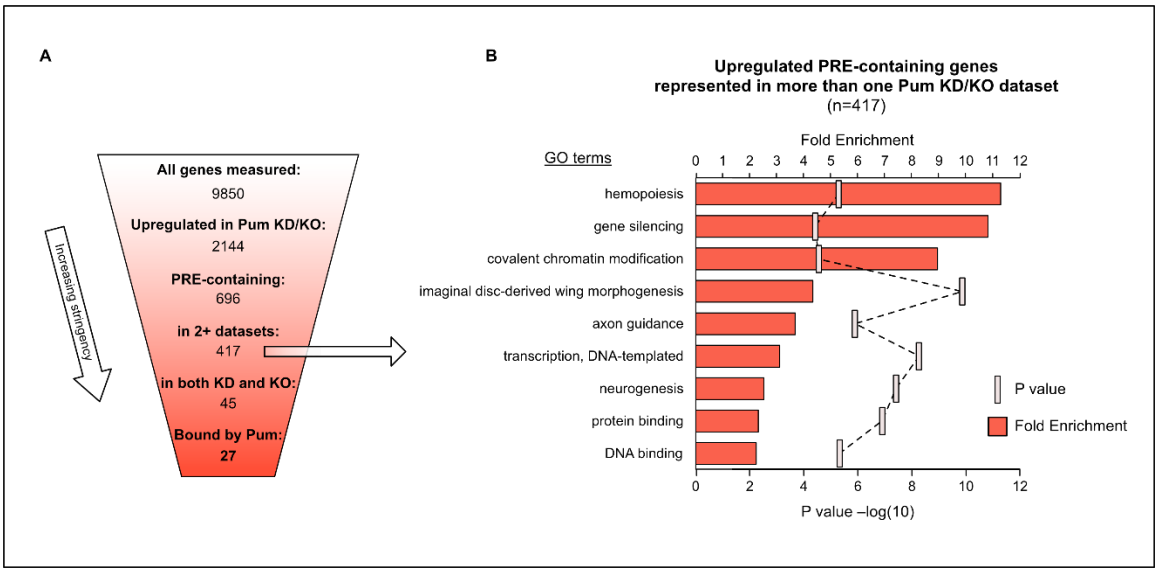


Figure 5.12: Genes upregulated in Pum KO/KD experiments. (A) A schematic representation of the filtering of genes from all datasets to obtain the highest confidence targets. (B) Gene ontology analysis of the 417 genes upregulated in two or more Pum KO/KD datasets. Fold enrichment is shown as full bars corresponding to the upper x axis. P value (as inverse log(10)) is shown as small bars and corresponds to the lower x axis. Data used to generate graph is reported in Appendix F.

Since most of these genes were from both Pum KO lines and we desired to obtain a broad sampling of Pum target genes under different conditions, we filtered this list of 417 genes further. Fifty-six non-overlapping genes were from the PAC-Seq and Click-Seq data. From those 56 genes, we isolated those that were also upregulated in either Pum KO line. After filtering, 45 high confidence genes remained. We also cross-referenced the remaining genes with the RIP-chip datasets to interrogate which of our 45

top Pum targets have evidence of being directly bound by Pum as well (Gerber, Luschnig et al. 2006, Laver, Li et al. 2015). This left us with 27 genes. These genes all contain PREs, are bound by Pum, and are upregulated when Pum is *both* depleted with RNAi and knocked out with CRISPR. These genes are listed in Table 5.2, along with their human homologs. Note that for the average fold change listed, we utilized the data for each experiment where the indicated gene was measured, not just the experiments where the gene is upregulated. This provides a more complete look at the expression level changes under Pum KO or KD.

Since most of these genes have mammalian homologs, we asked if there was evidence of the homologs being bound by Pum as well. We searched murine two datasets, one performed using iCLIP (individual nucleotide resolution cross linking immunoprecipitation) of Pum1 and Pum2 from mouse whole brain lysates (Zhang, Chen et al. 2017), and one using RIP-Chip from mouse testis (Chen, Zheng et al. 2012), and two datasets from human, both using RIP-Chip analysis from HeLa cells with PUM1 (Morris, Mukherjee et al. 2008), or both PUM1 and PUM2 (Galgano, Forrer et al. 2008). There were several genes bound by Pum homologs, indicating an evolutionarily conserved method of regulation (Table 5.2).

Gene name	FlyBase ID	# PRE	Pum KD/KO avg log(2) FC	*Drosophila* Binding data	Human homologs
CG43759	FBgn0264090	3	1.7	(2)	
bib	FBgn0000180	2	1.34	(2)	AQP(1-6,8,9)[1,3], MIP
ORMDL	FBgn0037110	2	1.13	(1), (2)	ORMDL(1-3)[3,4]
amx	FBgn0000077	2	1.03	(1)	TM2D3[3]
Jheh2	FBgn0034405	1	0.91	(1)	EPHX1
numb	FBgn0002973	1	0.87	(2)	NUMB[1,2], NUMBL
Best2	FBgn0035696	3	0.82	(1)	BEST(1-4)
MFS16	FBgn0034611	2	0.7	(1)	SLC37A(1,2,3)[3,4]
tkv	FBgn0003716	3	0.67	(1), (2)	ACVR1(A,B,C)[2,3], ACVRL1, BMPR1B, TGFBR1
CG34347	FBgn0085376	3	0.66	(1)	EPB41L4A
alpha4GT1	FBgn0031491	1	0.65	(1), (2)	A4GALT, A4GNT
Ppox	FBgn0020018	1	0.62	(1), (2)	PPOX
cher	FBgn0014141	2	0.61	(1)	FLNA, FLNB, FLNC
RnrS	FBgn0011704	1	0.6	(1), (2)	RRM2, RRM2B[3]
Coq7	FBgn0029502	1	0.6	(1), (2)	COQ7[4]
Raf	FBgn0003079	2	0.6	(1), (2)	BRAF, ARAF[2,3], RAF1
Kap-alpha1	FBgn0024889	1	0.59	(1)	KPNA(1-7)[3,4]
CG2200	FBgn0030447	2	0.59	(1), (2)	
CG4281	FBgn0025626	2	0.59	(1), (2)	
SCAR	FBgn0041781	2	0.56	(1), (2)	WASF(1,2,3)[3]
Ugalt	FBgn0024994	3	0.54	(1), (2)	SLC35A(1,2,3)[2,3,4]
CG18622	FBgn0038460	4	0.5	(1)	
tna	FBgn0026160	3	0.49	(1), (2)	ZMIZ1[2], ZMIZ2[1]
Cdc7	FBgn0028360	2	0.39	(1), (2)	CDC7[4]
Phm	FBgn0283509	2	0.36	(1), (2)	PAM[3]
Got2	FBgn0001125	1	0.34	(1)	GOT1, GOT2
CG7841	FBgn0036502	2	0.34	(1), (2)	MCRIP1, MCRIP2

Table 5.2: Top Pum target genes. Genes listed in this table are upregulated in two or more RNA-seq datasets, contain PREs, and are bound by Pum. Log(2) fold change (FC) shown is the mean FC from all datasets. *Drosophila* binding data from (1) (Gerber, Luschnig *et al*. 2006) and (2) (Laver, Li *et al*. 2015). Homologs identified in mouse or human binding data are indicated by superscripts as follows: homolog binding data from 1- (Chen, Zheng *et al*. 2012), 2- (Zhang, Chen *et al*. 2017), 3- (Galgano, Forrer *et al*. 2008), 4- (Morris, Mukherjee *et al*. 2008).

5.5. Summary and discussion

In this study we utilized PAC-Seq and Click-Seq to interrogate the transcriptome-wide changes when Pum is depleted. We also analyzed changes when Pum is knocked out using CRISPR-Cas9 gene editing. Our combined analyses of genes affected by Pum revealed over 2000 upregulated RNAs. 696 of these contain PREs, indicating they are direct targets. This data expands the known number of Pum controlled genes profoundly.

Using PAC-seq and Click-Seq, we also measured genome-wide changes with depletion of Not1 and Pop2. We found over 1,500 genes upregulated when either Not1 or Pop2 are knocked down, providing us with new insights into the regulatory network of the CNOT complex. Consistent with the CNOT complex's role in Pum-mediated decay, we found that a majority of Pum targets are also upregulated when Not1 and Pop2 are depleted. Of the 696 PRE-containing upregulated genes, 261 of these are also increased in abundance with Not1/Pop2 RNAi (odds ratio 5.8, 95% credible interval 4.6, 7.4).

Curiously, there were 33 genes that were upregulated in the Pum RNAi condition but either downregulated or not affected for Not1/Pop2 RNAi. 20 of these genes contain PREs. In the Click-Seq data alone, we found that 108 PRE-containing genes were upregulated when either Pop2 or Not1 were depleted, out of over 400 PRE-containing genes upregulated in the Pum-null lines. This percentage of co-regulated targets was lower than we expected. There are several possible explanations for this. First, some of the targets failed to meet our cutoffs for either biological or statistical significance in one of the datasets. Second, there could be factors with greater influence than Pum or CNOT acting on the transcripts. Third, the genes regulated by Pum but not CNOT might indicate a CNOT-independent mechanism of Pum-mediated decay. Finally, not all of Pum's targets are degraded. Pum proteins also act through translational inhibition which may not necessarily lead to a decrease in RNA levels (Chagnovich and Lehmann 2001, Chritton and Wickens 2011, Weidmann, Raynard et al. 2014).

Gene ontology of upregulated genes reaffirms the known links between Pum mutant phenotypical analyses to target RNAs. Wing morphogenesis involves genes from many developmental processes, Pum's most prominent historical role. Other established roles for Pum include axon guidance and neurogenesis, discussed briefly in Chapter 1.

Cancer phenotypes observed in mammals can be traced to dysregulation of the types of genes represented in these categories: transcription factors (Miles, Tschop et al. 2012, Oh, Shin et al. 2012, Naudin, Hattabi et al. 2017, Brocard, Khasnis et al. 2018), signaling molecules (Bermudez, Jouandin et al. 2011, Maurer, Tarkowski et al. 2011, Chen, Zheng et al. 2012), and cell cycle regulators (Belletti and Baldassarre 2015, Sun, Zhangyuan et al. 2017).

GO term analysis of the upregulated gene sets provide us with some new insight into processes Pum is involved in. Less is known about Pum's roles in gene silencing or chromatin modification. One of the genes in the gene silencing category is polycomb (Pc), a chromatin binding protein. Another is Pipsqueak (psq), a transcription factor that mediates the activity of polycomb group proteins and regulates chromatin silencing. We evaluated enhancer of polycomb, E(Pc), a subunit of the NuA4 histone acetyltransferase complex, and showed that it too is regulated by Pum (Chapter 4).

The genes populating the GO category hemopoiesis are transcription factors, signaling molecules, and other growth regulators. Raf, which we also evaluated in Chapter 4, is included in this set. E(Pc), significantly upregulated in both Pum KO lines (avg log(2) FC = 1.1), is also involved in hematopoiesis (Owusu-Ansah and Banerjee 2009). Additionally, human Pums have recently been linked to hematopoietic stem cell proliferation through the target FOXP1 (Naudin, Hattabi et al. 2017).

We also looked at the RNA levels for the targets we identified in Chapter 4. Chico was increased an average of 1.34-fold (log(2) 0.43) in the Pum KD conditions, but failed to meet the statistical significance cutoff ($p > 0.05$). Pten was significantly upregulated in both Pum KO lines. Raf was significantly increased in one KD condition and both KO lines, but both Pde11 and Nrv1 were not significantly upregulated in the Pum KD conditions. One explanation for the disparity in regulation seen between the RNAi experiments in Chapter 4 and RNA-seq analysis could be the duration of treatment. The experiments in Chapter 4 were optimized for Pum depletion and were 90-hour knockdowns. The RNA-seq experiments were 48 hours to limit the effects of CNOT depletion on cell viability.

We then examined some of the targets with documented Pum roles in *Drosophila*. Dap, the *Drosophila* homolog of CDKN1B (p27 protein), is significantly upregulated in both Pum KO lines. Egfr was demonstrated to be regulated by Pum (Kim,

Kim et al. 2012), and it is significantly upregulated for all Pum conditions. However, it does not appear in our top target list because it does not possess a perfect 8 nucleotide PRE, the criteria we used to define PRE-containing genes.

It is worth noting however, that the 8 nucleotide PRE with an adenine in position 8 is a stringent criteria and newer evidence in mammalian cells suggests the consensus sequence may contain either adenine or uracil in position 8 (UGUANAUA and UGUANAUU) (Chen, Zheng et al. 2012, Bohn, Van Etten et al. 2018, Wolfe, Schagat et al. 2020). Importantly, interactions with other proteins such as Nanos can alter the specificity of Pum and potentially increase the pool of targets (Weidmann, Qiu et al. 2016).

5.5.1. Biological functions of top targets in Notch signaling

Many of our top target genes for Pum have roles in signaling pathways, especially Notch signaling. Notch signaling is a significantly enriched GO term in all of our Pum depletion/knockout analyses. Here we summarize the known function of some of the top Pum targets involved in the Notch pathway.

The Notch signaling pathway is crucial during development for lateral inhibition in the neuroectoderm, and for overall homeostasis (Hori, Sen et al. 2013). Notch receptors are transmembrane proteins that undergo two internal cleavage events upon activation to release an intracellular domain (NICD) which translocates into the nucleus and activates the transcription of genes. The second cleavage event is catalyzed by a gamma secretase (De Strooper, Annaert et al. 1999, Yang, Zhou et al. 2019). *Drosophila* Amx is thought to function at the level of gamma secretase in the Notch pathway as a beta-amyloid peptide binding protein (Michellod and Randsholt 2008). Amx is also necessary for the proper localization of Notch receptors during early development (Das, Salazar et al. 2020). As gamma secretase is also responsible for cleaving amyloid precursor peptide (APP), the human homolog of Amx, TM2D3 has also been implicated in late-onset Alzheimer's disease (Zhang, Li et al. 2014, Jakobsdottir, van der Lee et al. 2016).

Lateral inhibition in the developing neuroectoderm results in a select number of cells destined to become neural, while the rest are forced toward an epidermal fate. The Bib (Big brain) gene encodes a transmembrane ion channel that acts synergistically with

Notch receptors to enforce the epidermal fate (Doherty 1997). The mechanism of this interaction is still unclear, though it is thought that Bib functions as an adhesive molecule, since Bib appears to have no channel activity, despite its homology to mammalian aquaporins (Tatsumi, Tsuji et al. 2009).

Drosophila Numb is a target that was upregulated in every RNA seq dataset for Pum, Not1 and Pop2 RNAi, and Pum KO. It is also bound by Pum (Laver, Li et al. 2015), and the mammalian homolog is bound by Pum1 and Pum2 (Chen, Zheng et al. 2012, Zhang, Chen et al. 2017). Numb is a membrane associated protein asymmetrically localized during the lateral inhibition process and contributes to cell fate determination (Uemura, Shepherd et al. 1989). Numb is a known antagonist of the Notch signaling pathway, interfering with the nuclear translocation of Su(H) and preventing expression of Notch target genes (Frise, Knoblich et al. 1996). Of note, Su(H) is also a Pum target identified in our data.

Apart from their roles in cell differentiation, Numb proteins also function as tumor suppressors by controlling p53 ubiquitination and degradation (Colaluca, Tosoni et al. 2008). Numb is the focus of many studies involving breast cancer research (Tosoni, Zecchini et al. 2015, Wang, Xiang et al. 2019), prostate cancer (Wang, Yang et al. 2021), colon cancer (Peng, Wang et al. 2019) and aging (George, Biressi et al. 2013, Zhao, Kiss et al. 2020).

Another significant Pum target involved in the Notch pathway is Tonalli (tna). Tna is a member of the protein inhibitor of activated STAT (PIAS) proteins, first identified as negative regulators of the Jak-STAT cytokine signaling pathway (Shuai and Liu 2005). Tna is a putative E3 ligase in the SUMOylation pathway and interacts with chromatin remodeling complexes (Monribot-Villanueva, Juarez-Uribe et al. 2013). Tna is necessary for development by controlling expression of *Drosophila* homeotic (Hox) genes (Rosales-Vega, Hernandez-Becerril et al. 2018). Both mammalian homologs of tna, Zmiz1 and Zmiz2 (previously Zimp10 and Zimp7) also function in the androgen receptor pathway (Beliakoff and Sun 2006). Additionally, Zmiz1 can interact directly with the Notch intracellular domain and affect the transcription of Notch genes, especially Myc (Pinnell, Yan et al. 2015). Zmiz2 has been shown to interact with β-catenin and co-regulate the transcription of genes in the Wnt signaling pathway (Lee, Zhu et al. 2013).

5.6. Materials and Methods

Generation of Pum KO cell lines is described in Materials & Methods for Chapter 4, Endogenous targets.

RNAi

RNAi was performed in Pum-Myc tagged *Drosophila* Dl1 cells. 20 µg of prepared dsRNA (Pum, Not1, Pop2, or LacZ NTC) was added to a drop of serum free complete media (Gibco) in the bottom of a 6-well dish. Cells were counted and diluted to a density of 3 x 10^6 cells/mL in 1 mL serum free media, then 1 mL was added to each well. Cells were incubated in dsRNA for 60 minutes, then 2 mL complete media containing 5% heat-inactivated FBS was added to each well. After 72 hours (PAC-Seq) or 48 hours (Click-Seq) cells were harvested. For the 72 hour experiment, viability was assessed using trypan blue staining prior to harvest. Then, the media was removed and cells were resuspended in 2 mL PBS. 1.5 mL of cells were taken for RNA, and the remaining 500 µL were processed for protein expression.

Pum KO cell lines

Pum KO cell lines and WT Dl1 cells were counted and diluted to a density of 1.5 x 10^6 cells/mL in complete media (Gibco). 2 mL were plated in 6-well plates in triplicate for each cell line. Cells were grown for 72 hours at 25°C then harvested and processed as for RNAi samples.

For verification of knockdown in tagged cell lines, harvested cells were resuspended in 2 mL PBS. 1.5 mL of the cells were taken for RNA, and the remaining 500 µL was pelleted and resuspended in 75 µL RIPA buffer with 2x protease inhibitors (Roche, Sigma). After a 10 minute incubation on ice, lysates were collected by centrifuging at 21,000 x g for 10 minutes and removing the supernatant. Protein concentration was measured using the Bio-Rad DC protein assay kit according to the manufacturer's directions.

20 µg of total protein extract was incubated with an equal volume of 2x SDS loading buffer at 85°C for ten minutes. Samples were loaded on 12% SDS-PAGE gels (laboratory-made) and run at 120 to 180 volts until dye front reached the bottom of the

gel. Gels were transferred to PVDF membranes (Millipore) for either 1.5 hours at 65 volts or overnight at 35 volts. Following transfer, membranes were either blocked in blotto for 1 hour or allowed to dry completely, then primary antibodies were applied for one hour at room temperature, or incubated overnight at 4°C. Membranes were washed three times for 10 minutes in blotto, then secondary antibodies were applied for 1 hour at room temperature. Membranes were washed three times for ten minutes each in blotto, then rinsed with 1x PBS before applying chemiluminescent substrate (Pierce- Thermo Fisher, or Immobilon- Millipore) Blots were imaged using the auto exposure setting on a ChemiDoc touch (BioRad).

Antibodies

All antibodies used in this study are listed in Appendix Mouse anti-Tubulin (CST, Cat# 3873) was diluted 1:1000 in blotto. Rabbit anti-Myc primary antibody (CST Cat# 2278S) was diluted 1:5000 in TBST. Goat anti-rabbit-HRP secondary antibody (CST, Cat# 7074P2 or Sigma Cat# AP187P) was diluted 1:5000 in blotto. Goat anti-mouse-HRP (Thermo Fisher Cat# 31430) was diluted 1:5000 in blotto. Total protein was quantified via detection by Sypro Ruby (Thermo Fisher) staining following manufacturer's directions.

RT-qPCR

RNA extraction was performed using the Maxwell RSC simply RNA Tissue RNA extraction kit (Promega) and concentration was determined using a NanoDrop spectrophotometer. A total of 10 µg RNA was taken from each sample for reverse transcription (RT) using GoScript (Promega) reverse transcriptase (5 µg for RT, 5 µg for noRT samples.) RT reactions were primed with random hexamers and carried out according to the manufacturer protocol. cDNA was then diluted with 100 µL water to a final concentration of approximately 41 ng/µL. Quantitative PCR (qPCR) was set up using Promega's Go-Taq qPCR Master Mix with 2 µL of each cDNA/noRT sample in a 20 µL reaction volume using either 100 nM or 200 nM final concentration qPCR primers for each gene. Reactions were performed on the Bio-Rad CFX96 instrument using the following cycling parameters: 3 min 98°C, [10 sec 95°C, 30 sec 62/63/64°C, 40 sec 72°C

+ image] x 39, 60°C-90°C melt curve + image. Individual annealing temps for each primer set and primer data are listed in Appendix D.

Data analysis for RT-qPCR

Fold change for each experiment was determined using the Pfaffl method where fold change is represented by the equation $FC = \frac{(E_{Target})^{\Delta Ct_{Target}}}{(E_{Ref})^{\Delta Ct_{Ref}}}$. E is defined as primer efficiency for each gene, and ∆Ct is Ct of the calibrator condition (LacZ) minus the Ct of the test condition (Pum KD or Not1KD). Reference genes were RPL32 for Pum KD and Act5c for Not1 KD. Unpaired Student's T tests with variances determined by F tests were performed using raw ∆Ct values for each gene.

PAC-seq library preparation

PAC-Seq was performed following the method described in (Routh, Ji et al. 2017, Elrod, Jaworski et al. 2019). RNA extraction was performed according to manufacturer's protocol using the Maxwell RSC simply RNA Tissue RNA extraction kit (Promega) and concentration was determined using the NanoDrop spectrophotometer. RNA samples were analyzed by Agilent Tapestation and resulting RIN (RNA integrity number) values were all 10. 2 µg of total cellular RNA was used to generate the PAC-Seq library. Reverse transcription was performed in a 20 µL reaction following manufacturer's protocol with Superscript III Reverse Transcriptase (Invitrogen) with the following components: up to 4 µg RNA, 1 µL of 5 mM AzVTP:dNTPs (at a ratio of 1:5 azido-nucleotides AzATP, AzCTP, AzGTP (AzVTP) to dNTPs), 1 µL 50 uM 3' Illumina_4N_21T primer (GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNN (21T)), 5x Superscript First Strand buffer, DTT, and RNase OUT (Invitrogen). RNA template was removed after cDNA synthesis by incubating for 20 minutes at 37°C with 10U RNase H (NEB). Azido-terminated cDNA was purified using the Zymo Clean and Concentrator kit and eluted in 10 µL of pH 7.2 50 mM HEPES buffer.

To form the click adapter-linked cDNA, all 10 µL of cDNA was incubated for 30 minutes at room temperature with 20 µL DMSO, 3 µL of 5 µM Click-Adapter (5' Hexynyl-12(N)AGATCGGaaGAGCGTCGTGTAGGGaaAGAGTGTAGATCTCGGTGGTCGCCGTATCATT) and 0.4 µL premixed 50mM vitamin C and 2 µL 10 mM Cu-TBTA (Lumiprobe).

The alkyne-azide cycloaddition of the adapter was catalyzed twice, then the click-linked cDNA was purified with a Zymo DNA column.

To anneal the remaining Illumina adapters (indexing primer CaaGCAGaaGA CGGCATACGAGATnnnnnnGTGACTGGAGTTCAGACGTGT, where nnnnnn is the index sequence, and universal primer aaTGATACGGCGACCACCGAG), PCR reactions were set up using 2.5 µL each of 5 µM primers, 5 µL click-ligated cDNA, 25 µL 2x OneTaq Standard Buffer Master Mix (NEB) in 50 µL total reactions. Optimized cycling parameters were 94° 4 min; 53° 30 sec; 68°10 min; [94° 30 sec, 53° 30 sec, 68° 2 min] x 20–22 cycles; 68° 5 min. Amplicons were size selected at 200 to 300 bp on a 2% agarose gel, then gel purified with the Zymo Gel DNA recovery kit. Final yields were determined with a QuBit fluorimeter.

Click-Seq library preparation

Click-Seq library preparation was performed following the methods developed in (Routh, Head et al. 2015). RNA extraction was performed using the Maxwell RSC simply RNA Tissue RNA extraction kit (Promega) and concentration was determined using the NanoDrop spectrophotometer. RNA samples were submitted for Agilent Tapestation analysis and RIN values were 8.7 and higher. Total RNA was poly(A) selected by oligo dT using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB Cat#E7490L). Library preparation then followed the steps described for PAC-Seq, with the exception of the primer sequence used during cDNA first strand synthesis: Illumina 6N p7 adapter (GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNN).

Sequencing was performed on pooled samples using manufacturer's standard operating procedures. Samples were sequenced with single end 75 bp reads on the NextSeq 500 (Illumina) with a high density v2.5 flow cell.

Data processing

The DPAC pipeline v1.10 (Routh 2019) was used to process the PAC-Seq results for differentially expressed genes (DE) in the resulting data. For the Click-Seq samples, sequencing adapters were trimmed and unique molecular identifiers (UMIs) annotated using fastp (version 0.14.1), then reads were aligned to the UCSC dm6 genome using hisat2 (version 2.1.0). The alignments were then deduplicated using UMI-tools (version 1.0.1) and differential expression was analyzed using DEseq2 (1.23.110)

and featureCounts (Rsubreads 1.30.9). Significance calling was based on adjusted p values (p-adj) and biological significance cutoff was 1.3-fold change.

Determination of PRE-containing genes

FlyBase release 6.38 February 2021 was used to download gene annotations for coding sequences (CDS), 3’ UTRs, 5’ UTRs, ncRNA, miRNA, and miscRNA. PRE processing was performed in R using the seqinR package. Data was then exported to Excel and gene matching based on FBgn, gene name, or gene locus was performed to assign PRE numbers to each gene from our RNA-seq datasets. Genes in RNA-seq data sets that could not be matched to any of the above mentioned FlyBase sequences were discarded. These included but were not limited to genes with withdrawn gene status, pseudogenes, some tRNAs, and mitochondrial genome genes.

Chapter 6: Conclusions

6.1. Project Summary

The scope of influence Pum proteins have over an entire organism is broad. Pum targets perform myriad functions in the cell. The ongoing investigations of Pum targets in both *Drosophila* and mammals have given us new insights into post transcriptional control mechanisms, not just for Pum but for the entire field.

6.1.1. Understanding Pum function

Pum proteins were identified decades ago and much work has been performed to understand their structure, function, and targets (Nusslein-Volhard, Frohnhofer et al. 1987). The initial characterization of the RNA binding domain led to a focus of interest on how Pum binds RNA and the importance of the RBD in the regulation of targets (Zamore, Williamson et al. 1997). Since the N-terminus of the protein lacks readily identifiable domains it had long been ignored, even as it was becoming clear that the RBD alone was not the only feature contributing to repression of targets (Wharton, Sonoda et al. 1998, Menon, Sanyal et al. 2004, Weston and Baines 2007, Muraro, Weston et al. 2008). The identification of the unique repression domains (RDs) of Pum and the development of cell-based assays to measure their activity have provided a large step forward in elucidating the mechanism of Pum repression (Weidmann and Goldstrohm 2012), yet several more questions remain.

Evidence in yeast had shown that PUF proteins cause mRNA decay and interact with deadenylases (Olivas and Parker 2000, Goldstrohm, Hook et al. 2006) and recent work on human Pum proteins indicated that this is a conserved mechanism (Van Etten, Schagat et al. 2012, Enwerem, Elrod et al. 2021). Finally, repression by Pum's RDs involves recruitment of the CNOT deadenylase complex in *Drosophila* (Arvola, Chang et al. 2020). From these studies a larger picture was emerging where Pum causes repression by interacting with the CNOT complex, not only through the RBD, but via its unstructured N-terminal RDs as well. Despite these advances, the specific interactions

with CNOT complex subunits were not identified. The work described in this thesis adds to our understanding of Pum function by addressing these gaps in knowledge.

Activity of RD3 depends on two phenylalanine residues.

The structure-function analysis in Chapter 2 identified several regions of conservation within the RD3 of Pum. Each one of these conserved regions (CRs) affects repression by RD3 when deleted. Without structural information for RD3, it is not possible to be certain if the effects are due to the loss of specific amino acids or an overall structural or shape change within the region when the CR is removed. Since RD3 is predicted to be unstructured, we assume that folding is not a major factor. Deletion of certain amino acids likely eliminates an interface used in protein-protein interactions.

The most important CR for RD3's activity is CR4. Deleting it alone causes a substantial decrease in activity, and when it is deleted in conjunction with CR1, all activity is lost. Within CR4 we identified two specific residues, F1033 and F1040, that are essential for repression. Additionally, we showed that CR1 and CR4 are both sufficient for RD3's repressive activity when multimerized. A tethered construct consisting of 4 repeats of CR1+4 can repress a reporter RNA as much as WT RD3, and a 5x CR1+4 construct has even greater activity than WT RD3. We then mutated the phenylalanine residues within this 5x CR1+4 construct to alanine and showed near complete loss of activity in their absence.

We further demonstrated the importance of these phenylalanine residues in chapter 4. We expressed constructs of RD3 with the RNA binding domain (RD3-RBD) and measured their activity on reporters carrying endogenous target 3'-UTRs. Wild type RD3-RBD was able to repress the reporters more than the RBD alone, but the double phenylalanine mutants were not.

RD3 interacts with multiple members of the CNOT complex.

We tested interactions between members of the *Drosophila* CNOT complex and RD3 in chapter 3. We found that Not1, Not2, and Not3 interact with RD3. We further isolated the specific region of each protein mediating those interactions. For Not1, the

interaction was localized to a region in the N-terminus of the protein where NOT1 is known to bind to TTP in humans (Tis11 in *Drosophila*), amino acids 908 to 1016. If Pum and Tis11 compete for binding of the CNOT complex, this could have implications for the RNA. A mutually exclusive binding arrangement could potentially affect the fate of a Pum or Tis11-bound target.

RD3 interacts with the C-terminal region of Not2 in the conserved NOT-box domain, amino acids 465 to 585. RD3 also interacts with the C-terminal region of Not3. Combined with the RD3 interaction with Not1-N, this discovery fits with the model of the CNOT complex as a modular assembly with flexible arrangements *in vivo*.

We then showed that the interaction between RD3 and Not1 is dependent on the phenylalanine residues in CR4. Mutating either F1033 or F1040 to alanine disrupts the interaction with Not1. The deletion of CR4 affects the interaction with Not2, but it is not enough to eliminate it. We screened deletions of all 5 identified CRs in RD3 for interactions with Not2 but were unable to identify the region of RD3 that Not2 is interacting with. We screened RD3 deletions of CRs for interactions with Not3-C and found that a double deletion of CR1 and CR4 disrupts the interaction.

We demonstrated that RD3 interacts with the CNOT complex in *Drosophila* cells using co-immunoprecipitation. The phenylalanine mutations in CR4 did not disrupt this association, which is likely reinforced through multiple contacts with both Not2 and Not3.

In addition to strengthening the interaction between Pum and the CNOT complex, multiple points of contact might modulate the inherent conformational flexibility of CNOT. This could, in turn, affect the deadenylase activity. It is worth noting that depletion of Not1, Not2, Not3, or Pop2 all affected RD3's repressive activity (Arvola, Chang et al. 2020).

6.1.2. Identifying Pum targets

RIP-Chip analyses have identified RNAs bound by Pum in *Drosophila* (Gerber, Luschnig et al. 2006, Laver, Li et al. 2015) but the capture and sequencing of bound RNA is not 100% efficient and targets of Pum cannot be identified by this manner alone. Our RNA sequencing data complements the existing data by adding a functional aspect. We identified thousands of genes in our Pum-null lines that were dysregulated in the absence of Pum. This not only reflects the biological necessity for Pum in the cell but

provides insight inside the mechanisms by which a cell can adjust to the absence of a crucial regulator. Using transient depletion of Pum has allowed us to identify some of the most responsive Pum targets. Our analysis has vastly expanded the known repertoire of Pum targets.

Gene ontology analysis of regulated Pum targets revealed facets of Pum control that have yet to be fully explored, such as gene silencing and hematopoiesis. Many of the genes regulated in *Drosophila* have mammalian homologs, and we noted conservation of Pum control in more than a dozen targets. While some of the targets we identified have well-known functions, less is known about their regulation. The effects of Pum regulating these genes could be a reduction in total protein levels, as is the case for Pde11 and Raf (Chapter 4). The discovery of Pum as a repressor of these genes will be especially important for research involving disease states that are sensitive to protein levels. Another outcome of Pum regulation could be to destabilize the pool of RNA and make the system more responsive to transcriptional input. Pum regulation is essential in systems where there is little transcriptional control such as neural synapses, or where transcription occurs too slowly, such as immediate signaling responses. This is pertinent as many of the genes we identified in Chapter 5 are involved in signaling and neural pathways.

6.2. Future directions

6.2.1. Interactions with the CNOT complex

We demonstrated that the physical interaction between RD3 and the CNOT complex is mediated by multiple protein contacts. RD3's CR4 is responsible for contacting Not1. We also identified contacts between RD3 and Not2 that do not depend on the CRs. *In vitro* biochemical assays have also affirmed the interaction between RD3 and the CNOT complex for *Drosophila* Pum and human CNOT proteins (Arvola, Chang et al. 2020, Enwerem, Elrod et al. 2021). However, these *in vitro* assays have thus far been limited in their ability to determine precise contacts for the individual proteins due to stability and expression challenges with single CNOT complex subunits. One pressing future goal is to purify a minimal peptide of Pum in complex with a CNOT module for

structural analysis via x-ray crystallography. Similar approaches have been used to co-crystallize Nanos peptides (Bhandari, Raisch et al. 2014) and a TTP peptide (Fabian, Frank et al. 2013) with Not1.

Mapping the protein contacts of the entire Pum protein will be needed to fully understand the mechanism of Pum repression. In the absence of structural information, the exploration of how Pum is contacting the CNOT complex will have to be completed through cell-based assays such as co-IPs and yeast two-hybrid assays. The regions of RD3 responsible for binding Not2 have yet to be identified. The Pum N-terminus (including RD1 and RD2) also contacts the CNOT complex (Arvola, Chang et al. 2020) and these contacts are unknown as well.

One of our more interesting discoveries was that the phenylalanine mutations that destroy the repressive activity of RD3 do not inhibit its association with the CNOT complex in *Drosophila* cells. This could be explored further in two ways. One, using *in vitro* deadenylation assays with recombinant wild type or mutant Pum and purified CNOT complex, such as those performed in Raisch et al. 2019, and two, *in vivo* utilizing a transcription shutoff approach with WT and mutant Pum. We would predict that while both WT and mutant Pum could bind and recruit CNOT, only the WT form would enhance deadenylation of target mRNAs.

This work focused on characterizing RD3 and its specific contacts to decay factors, especially the CNOT complex. We tested several other decay proteins for interactions but the results were negative. For future experiments, the Y2H system can be used to screen large libraries of proteins for *de novo* discovery of Pum partners. This analysis would not be limited to decay factors and could uncover links to other regulatory networks.

6.2.2. Pum function

Our finding in Chapter 3 that functionally inoperative RD3 mutations do not prohibit CNOT complex binding has implications for understanding overall Pum regulation. The repressive activity of Pum appears to be modulated by the presence of auto-regulatory domains, PCMa and PCMb (Weidmann and Goldstrohm 2012). These domains show a higher level of conservation than the RDs, which indicates their function is crucial to the activity of Pum. One of the possible mechanisms for inhibition of Pum

activity to occur could be through occlusion of a Pum-CNOT interaction by PCMb. This could inhibit complex binding or prevent potential stimulatory effects conferred by other Pum regions. Testing for interference of CNOT complex binding could be performed using Y2H assays or co-IPs.

Additionally, we identified a region of RD3, CR2, that also appears to have a regulatory function. Loss of activity resulting from deletions of the Q-P/N region, CR1, or CR3 is restored when CR2 is deleted in combination with them. Several observations were made that CR2 is a point of phosphorylation in Dmel2 cells (data not shown) but it is unclear under what circumstances the phosphorylation occurs and if it is biologically relevant. Vertebrate Pums have a highly-conserved insertion in this region that might alter or enhance the behavior of CR2. Future experiments could explore the contribution of CR2 to the regulation of Pum.

The N-terminus of *Drosophila* Pum contains other repression domains, RD1 and RD2. Motif analyses similar to the work described here for RD3 was performed for RD2 but remains incomplete (Joseph Buytendorp, unpublished data). The existing evidence that *Drosophila* RDs interact with the human CNOT complex indicates that each RD makes contacts with CNOT (Arvola, Chang et al. 2020). Future experiments can complete mapping of these contacts to the CNOT complex using mutational analysis and Y2H assays.

6.2.3. Pum targets

Our RNA-seq data uncovered many genes that are direct Pum targets. Follow up experiments would include verification of PRE-dependent decay using reporter assays, similar to the experiments carried out in Chapter 4. For select targets, CRISPR-Cas9 editing can be used to examine protein level changes with overexpression of Pum or Pum depletion. Additionally, the PRE of endogenous targets might be amenable to CRISPR editing, presenting the opportunity to engineer genes nonresponsive to Pum.

We used existing RIP-Chip data to supplement our sequencing results. However, many genes differentially expressed in our datasets do not have associated binding data. An ideal experiment would use RIP-seq or CLIP-seq with our Pum-Myc cell line to identify bound targets. This would provide a complementary dataset that could be

compared to the functionally regulated and PRE-containing gene sets to provide an extensive list of Pum bound and regulated mRNAs that contain a PRE.

Many genes in our sequencing datasets are downregulated upon Pum depletion. Although these can partially be attributed to secondary effects, that may not hold true for all of the decreased genes, some of which contain PREs. Recent evidence in human Pums suggests that repression or decay are not the only outcomes of Pum binding to a transcript. Indeed, Pum mediated activation of genes has been observed in mammals (Naudin, Hattabi et al. 2017, Bohn, Van Etten et al. 2018, Wolfe, Schagat et al. 2020). Pum acts to stabilize a certain subset of targets but the mechanism by which this occurs is entirely unknown. While there is no documented evidence of this phenomenon in *Drosophila* yet, the high degree of conservation between Pums in these species makes it an enticing objective for follow up.

Our sequencing data also revealed many genes that do not fit with our model of Pum-CNOT mediated repression. This could indicate alternative mechanisms of regulation. Two of these genes, *jheh2* and *Phm* appear in multiple datasets as being upregulated upon Pum depletion but downregulated for both Not1 and Pop2 depletion. These genes could serve as a baseline for exploring potential alternative decay mechanisms.

Our experiments in Chapter 4 touched upon a very pressing question: What makes a particular PRE more or less functional than another, and how is the outcome of Pum binding to a PRE determined?

One way that this may occur is through interactions with other proteins. The interaction with Nanos alters Pum's binding specificity and increase the pool of potential targets (Weidmann, Qiu et al. 2016). This has been documented in detail for cyclin B, where Pum alone has low affinity for the cyclin B RNA (Kadyrova, Habara et al. 2007, Weidmann, Qiu et al. 2016). Pum also cooperates with Nanos and Brat to regulate hunchback and paralytic RNA in early development (Arvola, Weidmann et al. 2017). For the regulation of paralytic, the requirement for Brat is cell type specific (Muraro, Weston et al. 2008). Pum also has context dependent interactions with CPEB where the number and proximity of PREs and CPEs to the polyadenylation sequence can determine the fate of the bound transcript. (Pique, Lopez et al. 2008, Campbell, Menichelli et al. 2012)

In addition to other proteins, Pum proteins potentially cooperate with the miRNA regulatory network (Galgano, Forrer et al. 2008). It has been suggested that Pum binding to a transcript can increase accessibility of miRNA seed sites, although the reverse might also occur (Kedde, van Kouwenhove et al. 2010, Miles, Tschop et al. 2012).

Large data sets such as those generated with RNA-seq present the opportunity for mining of motifs or patterns that could provide insight into other *cis* acting elements contributing to the fate of a target. This work has generated several large sets of data for such analyses. Follow up experiments could be applied to select targets or engineered reporters. Luciferase assays or overexpression/knockdown of regulators coupled to RT-qPCR can be used to measure regulatory effects.

While the overall consequence of Pum binding to a PRE is decay, the full extent of regulation is likely to be transcript specific, influenced by the full repertoire of *cis* elements, RNA structures, and transacting factors that act on that transcript. It is probable that there is no “one size fits all” model for the functionality and outcome of an individual PRE. This makes it even more important to understand how Pum functions, what other factors it associates with, and how it determines the fate of an RNA.

References

Ahringer, J., T. A. Rosenquist, D. N. Lawson and J. Kimble (1992). "The Caenorhabditis elegans sex determining gene fem-3 is regulated post-transcriptionally." EMBO J **11**(6): 2303-2310.

Arvola, R. M., C. T. Chang, J. P. Buytendorp, Y. Levdansky, E. Valkov, P. L. Freddolino and A. C. Goldstrohm (2020). "Unique repression domains of Pumilio utilize deadenylation and decapping factors to accelerate destruction of target mRNAs." Nucleic Acids Res **48**(4): 1843-1871.

Arvola, R. M., C. A. Weidmann, T. M. Tanaka Hall and A. C. Goldstrohm (2017). "Combinatorial control of messenger RNAs by Pumilio, Nanos and Brain Tumor Proteins." RNA Biol **14**(11): 1445-1456.

Asaoka-Taguchi, M., M. Yamada, A. Nakamura, K. Hanyu and S. Kobayashi (1999). "Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos." Nat Cell Biol **1**(7): 431-437.

Bachorik, J. L. and J. Kimble (2005). "Redundant control of the Caenorhabditis elegans sperm/oocyte switch by PUF-8 and FBF-1, two distinct PUF RNA-binding proteins." Proc Natl Acad Sci U S A **102**(31): 10893-10897.

Baek, K. H., J. R. Fabian, F. Sprenger, D. K. Morrison and L. Ambrosio (1996). "The activity of D-raf in torso signal transduction is altered by serine substitution, N-terminal deletion, and membrane targeting." Dev Biol **175**(2): 191-204.

Baines, R. A. (2005). "Neuronal homeostasis through translational control." Molecular Neurobiology **32**(2): 113-121.

Banerjee, A. K. (1980). "5'-terminal cap structure in eucaryotic messenger ribonucleic acids." Microbiological Reviews **44**(2): 175-205.

Baralle, F. E. and J. Giudice (2017). "Alternative splicing as a regulator of development and tissue identity." Nat Rev Mol Cell Biol **18**(7): 437-451.

Barbee, S. A., P. S. Estes, A. M. Cziko, J. Hillebrand, R. A. Luedeman, J. M. Coller, N. Johnson, I. C. Howlett, C. Geng, R. Ueda, A. H. Brand, S. F. Newbury, J. E. Wilhelm, R. B. Levine, A. Nakamura, R. Parker and M. Ramaswami (2006). "Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic P bodies." Neuron **52**(6): 997-1009.

Barker, D. D., C. Wang, J. Moore, L. K. Dickinson and R. Lehmann (1992). "Pumilio is essential for function but not for distribution of the Drosophila abdominal determinant Nanos." Genes Dev **6**: 2312-2326.

Bartel, D. P. (2009). "MicroRNAs: target recognition and regulatory functions." Cell **136**(2): 215-233.

Bashirullah, A., R. L. Cooperstock and H. D. Lipshitz (1998). "RNA localization in development." Annu Rev Biochem **67**: 335-394.

Basquin, J., V. V. Roudko, M. Rode, C. Basquin, B. Seraphin and E. Conti (2012). "Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction." Mol Cell **48**(2): 207-218.

Bassett, A. R., C. Tibbit, C. P. Ponting and J. L. Liu (2014). "Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9." Biol Open **3**(1): 42-49.

Bawankar, P., B. Loh, L. Wohlbold, S. Schmidt and E. Izaurralde (2013). "NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain." RNA Biol **10**(2): 228-244.
Behm-Ansmant, I., J. Rehwinkel, T. Doerks, A. Stark, P. Bork and E. Izaurralde (2006). "mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes." Genes Dev **20**(14): 1885-1898.
Beliakoff, J. and Z. Sun (2006). "Zimp7 and Zimp10, two novel PIAS-like proteins, function as androgen receptor coregulators." Nucl Recept Signal **4**: e017.
Belletti, B. and G. Baldassarre (2015). "Roles of CDKN1B in cancer?" Aging (Albany NY) **7**(8): 529-530.
Benoit, P., C. Papin, J. E. Kwak, M. Wickens and M. Simonelig (2008). "PAP- and GLD-2-type poly(A) polymerases are required sequentially in cytoplasmic polyadenylation and oogenesis in Drosophila." Development **135**(11): 1969-1979.
Bentley, D. (2002). "The mRNA assembly line: transcription and processing machines in the same factory." Current Opinion in Cell Biology **14**(3): 336-342.
Bermudez, O., P. Jouandin, J. Rottier, C. Bourcier, G. Pages and C. Gimond (2011). "Post-transcriptional regulation of the DUSP6/MKP-3 phosphatase by MEK/ERK signaling and hypoxia." J Cell Physiol **226**(1): 276-284.
Bhandari, D., T. Raisch, O. Weichenrieder, S. Jonas and E. Izaurralde (2014). "Structural basis for the Nanos-mediated recruitment of the CCR4-NOT complex and translational repression." Genes Dev **28**(8): 888-901.
Blewett, N. H. and A. C. Goldstrohm (2012). "A eukaryotic translation initiation factor 4E-binding protein promotes mRNA decapping and is required for PUF repression." Mol Cell Biol **32**(20): 4181-4194.
Blower, M. D. (2013). "Molecular insights into intracellular RNA localization." Int Rev Cell Mol Biol **302**: 1-39.
Bodenmiller, B., J. Malmstrom, B. Gerrits, D. Campbell, H. Lam, A. Schmidt, O. Rinner, L. N. Mueller, P. T. Shannon, P. G. Pedrioli, C. Panse, H. K. Lee, R. Schlapbach and R. Aebersold (2007). "PhosphoPep--a phosphoproteome resource for systems biology research in Drosophila Kc167 cells." Mol Syst Biol **3**: 139.
Bohn, J. A., J. L. Van Etten, T. L. Schagat, B. M. Bowman, R. C. McEachin, P. L. Freddolino and A. C. Goldstrohm (2018). "Identification of diverse target RNAs that are functionally regulated by human Pumilio proteins." Nucleic Acids Res **46**(1): 362-386.
Böhni, R., J. Riesgo-Escovar, S. Oldham, W. Brogiolo, H. Stocker, B. F. Andruss, K. Beckingham and E. Hafen (1999). "Autonomous Control of Cell and Organ Size by CHICO, a Drosophila Homolog of Vertebrate IRS1–4." Cell **97**(7): 865-875.
Boland, A., Y. Chen, T. Raisch, S. Jonas, D. Kuzuoglu-Ozturk, L. Wohlbold, O. Weichenrieder and E. Izaurralde (2013). "Structure and assembly of the NOT module of the human CCR4-NOT complex." Nat Struct Mol Biol **20**(11): 1289-1297.
Borah, S., A. C. Wong and J. A. Steitz (2009). "Drosophila hnRNP A1 homologs Hrp36/Hrp38 enhance U2-type versus U12-type splicing to regulate alternative splicing of the prospero twintron." Proc Natl Acad Sci U S A **106**(8): 2577-2582.
Braun, J. E., E. Huntzinger, M. Fauser and E. Izaurralde (2011). "GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets." Mol Cell **44**(1): 120-133.
Braun, J. E., V. Truffault, A. Boland, E. Huntzinger, C. T. Chang, G. Haas, O. Weichenrieder, M. Coles and E. Izaurralde (2012). "A direct interaction between DCP1

and XRN1 couples mRNA decapping to 5' exonucleolytic degradation." Nat Struct Mol Biol **19**(12): 1324-1331.
Brengues, M., D. Teixeira and R. Parker (2005). "Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies." Science **310**(5747): 486-489.
Brocard, M., S. Khasnis, C. D. Wood, C. Shannon-Lowe and M. J. West (2018). "Pumilio directs deadenylation-associated translational repression of the cyclin-dependent kinase 1 activator RGC-32." Nucleic Acids Res.
Brose, M. S., P. Volpe, M. Feldman, M. Kumar, I. Rishi, R. Gerrero, E. Einhorn, M. Herlyn, J. Minna, A. Nicholson, J. A. Roth, S. M. Albelda, H. Davies, C. Cox, G. Brignell, P. Stephens, P. A. Futreal, R. Wooster, M. R. Stratton and B. L. Weber (2002). "BRAF and RAS mutations in human lung cancer and melanoma." Cancer Res **62**(23): 6997-7000.
Campbell, Z. T., E. Menichelli, K. Friend, J. Wu, J. Kimble, J. R. Williamson and M. Wickens (2012). "Identification of a conserved interface between PUF and CPEB proteins." J Biol Chem **287**(22): 18854-18862.
Carreira-Rosario, A., V. Bhargava, J. Hillebrand, R. K. Kollipara, M. Ramaswami and M. Buszczak (2016). "Repression of Pumilio Protein Expression by Rbfox1 Promotes Germ Cell Differentiation." Dev Cell **36**(5): 562-571.
Chagnovich, D. and R. Lehmann (2001). "Poly(A)-independent regulation of maternal hunchback translation in the Drosophila embryo." Proc Natl Acad Sci U S A **98**(20): 11359-11364.
Chang, C. T., S. Muthukumar, R. Weber, Y. Levdansky, Y. Chen, D. Bhandari, C. Igreja, L. Wohlbold, E. Valkov and E. Izaurralde (2019). "A low-complexity region in human XRN1 directly recruits deadenylation and decapping factors in 5'-3' messenger RNA decay." Nucleic Acids Res **47**(17): 9282-9295.
Chen, D., W. Zheng, A. Lin, K. Uyhazi, H. Zhao and H. Lin (2012). "Pumilio 1 suppresses multiple activators of p53 to safeguard spermatogenesis." Curr Biol **22**(5): 420-425.
Chen, G., W. Li, Q. S. Zhang, M. Regulski, N. Sinha, J. Barditch, T. Tully, A. R. Krainer, M. Q. Zhang and J. Dubnau (2008). "Identification of synaptic targets of Drosophila pumilio." PLoS Comput Biol **4**(2): e1000026.
Chen, Y., C. Bao, X. Zhang, X. Lin, H. Huang and Z. Wang (2019). "Long non-coding RNA HCG11 modulates glioma progression through cooperating with miR-496/CPEB3 axis." Cell Prolif **52**(5): e12615.
Chen, Y., A. Boland, D. Kuzuoglu-Ozturk, P. Bawankar, B. Loh, C. T. Chang, O. Weichenrieder and E. Izaurralde (2014). "A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing." Mol Cell **54**(5): 737-750.
Cho, E. J., T. Takagi, C. R. Moore and S. Buratowski (1997). "mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain." Genes Dev **11**(24): 3319-3326.
Choudhury, S. R., A. K. Singh, T. McLeod, M. Blanchette, B. Jang, P. Badenhorst, A. Kanhere and S. Brogna (2016). "Exon junction complex proteins bind nascent transcripts independently of pre-mRNA splicing in Drosophila melanogaster." Elife **5**.
Chritton, J. J. and M. Wickens (2011). "A role for the poly(A)-binding protein Pab1p in PUF protein-mediated repression." J Biol Chem **286**(38): 33268-33278.

Colaluca, I. N., D. Tosoni, P. Nuciforo, F. Senic-Matuglia, V. Galimberti, G. Viale, S. Pece and P. P. Di Fiore (2008). "NUMB controls p53 tumour suppressor activity." Nature **451**(7174): 76-80.
Crittenden, S. L., D. S. Bernstein, J. L. Bachorik, B. E. Thompson, M. Gallegos, A. G. Petcherski, G. Moulder, R. Barstead, M. Wickens and J. Kimble (2002). "A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans." Nature **417**(6889): 660-663.
D'Amico, D., A. Mottis, F. Potenza, V. Sorrentino, H. Li, M. Romani, V. Lemos, K. Schoonjans, N. Zamboni, G. Knott, B. L. Schneider and J. Auwerx (2019). "The RNA-Binding Protein PUM2 Impairs Mitochondrial Dynamics and Mitophagy During Aging." Mol Cell **73**(4): 775-787 e710.
Dabrowski, M., Z. Bukowy-Bieryllo and E. Zietkiewicz (2015). "Translational readthrough potential of natural termination codons in eucaryotes--The impact of RNA sequence." RNA Biol **12**(9): 950-958.
Dantonel, J. C., K. G. Murthy, J. L. Manley and L. Tora (1997). "Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA." Nature **389**(6649): 399-402.
Das, P., J. L. Salazar, D. Li-Kroeger, S. Yamamoto, M. Nakamura, T. Sasamura, M. Inaki, W. Masuda, M. Kitagawa, T. Yamakawa and K. Matsuno (2020). "Maternal almondex, a neurogenic gene, is required for proper subcellular Notch distribution in early Drosophila embryogenesis." Dev Growth Differ **62**(1): 80-93.
Day, J. P., J. A. Dow, M. D. Houslay and S. A. Davies (2005). "Cyclic nucleotide phosphodiesterases in Drosophila melanogaster." Biochem J **388**(Pt 1): 333-342.
De Strooper, B., W. Annaert, P. Cupers, P. Saftig, K. Craessaerts, J. S. Mumm, E. H. Schroeter, V. Schrijvers, M. S. Wolfe, W. J. Ray, A. Goate and R. Kopan (1999). "A presenilin-1-dependent gamma-secretase-like protease mediates release of Notch intracellular domain." Nature **398**(6727): 518-522.
Dever, T. E., J. D. Dinman and R. Green (2018). "Translation Elongation and Recoding in Eukaryotes." Cold Spring Harb Perspect Biol **10**(8).
Dieci, G., G. Fiorino, M. Castelnuovo, M. Teichmann and A. Pagano (2007). "The expanding RNA polymerase III transcriptome." Trends Genet **23**(12): 614-622.
Doherty, D., Jan, L. Y., Jan, Y. N. (1997). "The Drosophila neurogenic gene big brain, which encodes a membrane-associated protein, acts cell autonomously and can act synergistically with Notch and Delta." Development **124**: 3881-3893.
Doidge, R., S. Mittal, A. Aslam and G. S. Winkler (2012). "The anti-proliferative activity of BTG/TOB proteins is mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) deadenylase subunits of the Ccr4-not complex." PLoS One **7**(12): e51331.
Doyon, Y. and J. Cote (2004). "The highly conserved and multifunctional NuA4 HAT complex." Curr Opin Genet Dev **14**(2): 147-154.
Dubnau, J., A. S. Chiang, L. Grady, J. Barditch, S. Gossweiler, J. McNeil, P. Smith, F. Buldoc, R. Scott, U. Certa, C. Broger and T. Tully (2003). "The staufen/pumilio pathway is involved in Drosophila long-term memory." Current Biology **13**(4): 286-296.
Dunn, J. G., C. K. Foo, N. G. Belletier, E. R. Gavis and J. S. Weissman (2013). "Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster." Elife **2**: e01179.
Edwalds-Gilbert, G., K. L. Veraldi and C. Milcarek (1997). "Alternative poly(A) site selection in complex transcription units: means to an end?" Nucleic Acids Res **25**(13): 2547-2561.

Eichhorn, S. W., A. O. Subtelny, I. Kronja, J. C. Kwasnieski, T. L. Orr-Weaver and D. P. Bartel (2016). "mRNA poly(A)-tail changes specified by deadenylation broadly reshape translation in Drosophila oocytes and early embryos." Elife **5**.
Elbashir, S. M., J. Martinez, A. Patkaniowska, W. Lendeckel and T. Tuschl (2001). "Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate." EMBO J **20**(23): 6877-6888.
Elgin, S. C. and G. Reuter (2013). "Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila." Cold Spring Harb Perspect Biol **5**(8): a017780.
Elrod, N. D., E. A. Jaworski, P. Ji, E. J. Wagner and A. Routh (2019). "Development of Poly(A)-ClickSeq as a tool enabling simultaneous genome-wide poly(A)-site identification and differential expression analysis." Methods **155**: 20-29.
Enwerem, III, N. D. Elrod, C. T. Chang, A. Lin, P. Ji, J. A. Bohn, Y. Levdansky, E. J. Wagner, E. Valkov and A. C. Goldstrohm (2021). "Human Pumilio proteins directly bind the CCR4-NOT deadenylase complex to regulate the transcriptome." RNA **27**(4): 445-464.
Fabian, M. R., M. K. Cieplak, F. Frank, M. Morita, J. Green, T. Srikumar, B. Nagar, T. Yamamoto, B. Raught, T. F. Duchaine and N. Sonenberg (2011). "miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT." Nat Struct Mol Biol **18**(11): 1211-1217.
Fabian, M. R., F. Frank, C. Rouya, N. Siddiqui, W. S. Lai, A. Karetnikov, P. J. Blackshear, B. Nagar and N. Sonenberg (2013). "Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin." Nat Struct Mol Biol **20**(6): 735-739.
Fabian, M. R. and N. Sonenberg (2012). "The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC." Nat Struct Mol Biol **19**(6): 586-593.
Fiedler, M., M. Graeb, J. Mieszczanek, T. J. Rutherford, C. M. Johnson and M. Bienz (2015). "An ancient Pygo-dependent Wnt enhanceosome integrated by Chip/LDB-SSDP." Elife **4**.
Fiore, R., S. Khudayberdiev, M. Christensen, G. Siegel, S. W. Flavell, T. K. Kim, M. E. Greenberg and G. Schratt (2009). "Mef2-mediated transcription of the miR379-410 cluster regulates activity-dependent dendritogenesis by fine-tuning Pumilio2 protein levels." EMBO J **28**(6): 697-710.
Fiore, R., M. Rajman, C. Schwale, S. Bicker, A. Antoniou, C. Bruehl, A. Draguhn and G. Schratt (2014). "MiR-134-dependent regulation of Pumilio-2 is necessary for homeostatic synaptic depression." EMBO J **33**(19): 2231-2246.
Forbes, A. and R. Lehmann (1998). "Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells." Development **125**(4): 679-690.
Friedman, A. A., G. Tucker, R. Singh, D. Yan, A. Vinayagam, Y. Hu, R. Binari, P. Hong, X. Sun, M. Porto, S. Pacifico, T. Murali, R. L. Finley, Jr., J. M. Asara, B. Berger and N. Perrimon (2011). "Proteomic and functional genomic landscape of receptor tyrosine kinase and ras to extracellular signal-regulated kinase signaling." Sci Signal **4**(196): rs10.
Friend, K., Z. T. Campbell, A. Cooke, P. Kroll-Conner, M. P. Wickens and J. Kimble (2012). "A conserved PUF-Ago-eEF1A complex attenuates translation elongation." Nat Struct Mol Biol **19**(2): 176-183.

Frise, E., J. A. Knoblich, S. Younger-Shepherd, L. Y. Jan and Y. N. Jan (1996). "The Drosophila Numb protein inhibits signaling of the Notch receptor during cell-cell interaction in sensory organ lineage." Proc Natl Acad Sci U S A **93**(21): 11925-11932.
Fromm, S. A., V. Truffault, J. Kamenz, J. E. Braun, N. A. Hoffmann, E. Izaurralde and R. Sprangers (2012). "The structural basis of Edc3- and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex." EMBO J **31**(2): 279-290.
Funakoshi, Y., Y. Doi, N. Hosoda, N. Uchida, M. Osawa, I. Shimada, M. Tsujimoto, T. Suzuki, T. Katada and S. Hoshino (2007). "Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases." Genes Dev **21**(23): 3135-3148.
Galgano, A., M. Forrer, L. Jaskiewicz, A. Kanitz, M. Zavolan and A. P. Gerber (2008). "Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system." PLoS One **3**(9): e3164.
Ganser, L. R., M. L. Kelly, D. Herschlag and H. M. Al-Hashimi (2019). "The roles of structural dynamics in the cellular functions of RNAs." Nat Rev Mol Cell Biol **20**(8): 474-489.
Garces, R. G., W. Gillon and E. F. Pai (2007). "Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties." Protein Sci **16**(2): 176-188.
Garneau, N. L., J. Wilusz and C. J. Wilusz (2007). "The highways and byways of mRNA decay." Nat Rev Mol Cell Biol **8**(2): 113-126.
Gebert, L. F. R. and I. J. MacRae (2019). "Regulation of microRNA function in animals." Nat Rev Mol Cell Biol **20**(1): 21-37.
Gennarino, V. A., E. E. Palmer, L. M. McDonell, L. Wang, C. J. Adamski, A. Koire, L. See, C.-A. Chen, C. P. Schaaf, J. A. Rosenfeld, J. A. Panzer, U. Moog, S. Hao, A. Bye, E. P. Kirk, P. Stankiewicz, A. M. Breman, A. McBride, T. Kandula, H. A. Dubbs, R. Macintosh, M. Cardamone, Y. Zhu, K. Ying, K.-R. Dias, M. T. Cho, L. B. Henderson, B. Baskin, P. Morris, J. Tao, M. J. Cowley, M. E. Dinger, T. Roscioli, O. Caluseriu, O. Suchowersky, R. K. Sachdev, O. Lichtarge, J. Tang, K. M. Boycott, J. L. Holder and H. Y. Zoghbi (2018). "A Mild PUM1 Mutation Is Associated with Adult-Onset Ataxia, whereas Haploinsufficiency Causes Developmental Delay and Seizures." Cell **172**(5): 924-936.e911.
Gennarino, V. A., R. K. Singh, J. J. White, A. De Maio, K. Han, J. Y. Kim, P. Jafar-Nejad, A. di Ronza, H. Kang, L. S. Sayegh, T. A. Cooper, H. T. Orr, R. V. Sillitoe and H. Y. Zoghbi (2015). "Pumilio1 haploinsufficiency leads to SCA1-like neurodegeneration by increasing wild-type Ataxin1 levels." Cell **160**(6): 1087-1098.
George, R. M., S. Biressi, B. J. Beres, E. Rogers, A. K. Mulia, R. E. Allen, A. Rawls, T. A. Rando and J. Wilson-Rawls (2013). "Numb-deficient satellite cells have regeneration and proliferation defects." Proc Natl Acad Sci U S A **110**(46): 18549-18554.
Gerber, A. P., S. Luschnig, M. A. Krasnow, P. O. Brown and D. Herschlag (2006). "Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in Drosophila melanogaster." Proc Natl Acad Sci U S A **103**(12): 4487-4492.
Goberdhan, D. C., N. Paricio, E. C. Goodman, M. Mlodzik and C. Wilson (1999). "Drosophila tumor suppressor PTEN controls cell size and number by antagonizing the Chico/PI3-kinase signaling pathway." Genes Dev **13**(24): 3244-3258.
Godwin, A. R., S. Kojima, C. B. Green and J. Wilusz (2013). "Kiss your tail goodbye: the role of PARN, Nocturnin, and Angel deadenylases in mRNA biology." Biochim Biophys Acta **1829**(6-7): 571-579.

Goldstrohm, A. C., T. M. T. Hall and K. M. McKenney (2018). "Post-transcriptional Regulatory Functions of Mammalian Pumilio Proteins." Trends Genet **34**(12): 972-990.
Goldstrohm, A. C., B. A. Hook, D. J. Seay and M. Wickens (2006). "PUF proteins bind Pop2p to regulate messenger RNAs." Nat Struct Mol Biol **13**(6): 533-539.
Goldstrohm, A. C., D. J. Seay, B. A. Hook and M. Wickens (2007). "PUF protein-mediated deadenylation is catalyzed by Ccr4p." J Biol Chem **282**(1): 109-114.
Gonatopoulos-Pournatzis, T. and V. H. Cowling (2014). "Cap-binding complex (CBC)." Biochem J **457**(2): 231-242.
Hanson, G. and J. Coller (2018). "Codon optimality, bias and usage in translation and mRNA decay." Nat Rev Mol Cell Biol **19**(1): 20-30.
Hara, M., S. Lourido, B. Petrova, H. J. Lou, J. R. Von Stetina, H. Kashevsky, B. E. Turk and T. L. Orr-Weaver (2018). "Identification of PNG kinase substrates uncovers interactions with the translational repressor TRAL in the oocyte-to-embryo transition." Elife **7**.
Hellen, C. U. T. (2018). "Translation Termination and Ribosome Recycling in Eukaryotes." Cold Spring Harb Perspect Biol **10**(10).
Hilgers, V., S. B. Lemke and M. Levine (2012). "ELAV mediates 3' UTR extension in the Drosophila nervous system." Genes Dev **26**(20): 2259-2264.
Hinnebusch, A. G., I. P. Ivanov and N. Sonenberg (2016). "Translational control by 5'-untranslated regions of eukaryotic mRNAs." Science **352**(6292): 1413-1416.
Hori, K., A. Sen and S. Artavanis-Tsakonas (2013). "Notch signaling at a glance." J Cell Sci **126**(Pt 10): 2135-2140.
Horvath, A., S. Boikos, C. Giatzakis, A. Robinson-White, L. Groussin, K. J. Griffin, E. Stein, E. Levine, G. Delimpasi, H. P. Hsiao, M. Keil, S. Heyerdahl, L. Matyakhina, R. Libe, A. Fratticci, L. S. Kirschner, K. Cramer, R. C. Gaillard, X. Bertagna, J. A. Carney, J. Bertherat, I. Bossis and C. A. Stratakis (2006). "A genome-wide scan identifies mutations in the gene encoding phosphodiesterase 11A4 (PDE11A) in individuals with adrenocortical hyperplasia." Nat Genet **38**(7): 794-800.
Hosoda, N., Y. Funakoshi, M. Hirasawa, R. Yamagishi, Y. Asano, R. Miyagawa, K. Ogami, M. Tsujimoto and S. Hoshino (2011). "Anti-proliferative protein Tob negatively regulates CPEB3 target by recruiting Caf1 deadenylase." EMBO J **30**(7): 1311-1323.
Huang, Y. H., C. C. Wu, C. K. Chou and C. Y. Huang (2011). "A translational regulator, PUM2, promotes both protein stability and kinase activity of Aurora-A." PLoS One **6**(5): e19718.
Hutvagner, G. and P. D. Zamore (2002). "A microRNA in a multiple-turnover RNAi enzyme complex." Science **297**(5589): 2056-2060.
Iglesias, N. and F. Stutz (2008). "Regulation of mRNP dynamics along the export pathway." FEBS Lett **582**(14): 1987-1996.
Jackson, R. J., C. U. Hellen and T. V. Pestova (2010). "The mechanism of eukaryotic translation initiation and principles of its regulation." Nat Rev Mol Cell Biol **11**(2): 113-127.
Jakobsdottir, J., S. J. van der Lee, J. C. Bis, V. Chouraki, D. Li-Kroeger, S. Yamamoto, M. L. Grove, A. Naj, M. Vronskaya, J. L. Salazar, A. L. DeStefano, J. A. Brody, A. V. Smith, N. Amin, R. Sims, C. A. Ibrahim-Verbaas, S. H. Choi, C. L. Satizabal, O. L. Lopez, A. Beiser, M. A. Ikram, M. E. Garcia, C. Hayward, T. V. Varga, S. Ripatti, P. W. Franks, G. Hallmans, O. Rolandsson, J. H. Jansson, D. J. Porteous, V. Salomaa, G. Eiriksdottir, K. M. Rice, H. J. Bellen, D. Levy, A. G. Uitterlinden, V. Emilsson, J. I. Rotter, T. Aspelund, H. Cohorts for, c. Aging Research in Genomic Epidemiology, C.

Alzheimer's Disease Genetic, Genetic, c. Environmental Risk in Alzheimer's Disease, C. J. O'Donnell, A. L. Fitzpatrick, L. J. Launer, A. Hofman, L. S. Wang, J. Williams, G. D. Schellenberg, E. Boerwinkle, B. M. Psaty, S. Seshadri, J. M. Shulman, V. Gudnason and C. M. van Duijn (2016). "Rare Functional Variant in TM2D3 is Associated with Late-Onset Alzheimer's Disease." PLoS Genet **12**(10): e1006327.
Ji, Y. and A. V. Tulin (2016). "Poly(ADP-Ribosyl)ation of hnRNP A1 Protein Controls Translational Repression in Drosophila." Mol Cell Biol **36**(19): 2476-2486.
Johannes, B. and A. Preiss (2002). "Wing vein formation in Drosophila melanogaster: Hairless is involved in the cross-talk between Notch and EGF signaling pathways." Mechanisms of Development **115**(1-2): 3-14.
Joly, W., A. Chartier, P. Rojas-Rios, I. Busseau and M. Simonelig (2013). "The CCR4 deadenylase acts with Nanos and Pumilio in the fine-tuning of Mei-P26 expression to promote germline stem cell self-renewal." Stem Cell Reports **1**(5): 411-424.
Jungreis, I., M. F. Lin, R. Spokony, C. S. Chan, N. Negre, A. Victorsen, K. P. White and M. Kellis (2011). "Evidence of abundant stop codon readthrough in Drosophila and other metazoa." Genome Res **21**(12): 2096-2113.
Kadyrova, L. Y., Y. Habara, T. H. Lee and R. P. Wharton (2007). "Translational control of maternal Cyclin B mRNA by Nanos in the Drosophila germline." Development **134**(8): 1519-1527.
Kashima, I., M. Takahashi, Y. Hashimoto, E. Sakota, Y. Nakamura and T. Inada (2014). "A functional involvement of ABCE1, eukaryotic ribosome recycling factor, in nonstop mRNA decay in Drosophila melanogaster cells." Biochimie **106**: 10-16.
Kaufmann, I., G. Martin, A. Friedlein, H. Langen and W. Keller (2004). "Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase." EMBO J **23**(3): 616-626.
Kedde, M., M. van Kouwenhove, W. Zwart, J. A. Oude Vrielink, R. Elkon and R. Agami (2010). "A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility." Nat Cell Biol **12**(10): 1014-1020.
Kedersha, N., S. Chen, N. Gilks, W. Li, I. J. Miller, J. Stahl and P. Anderson (2002). "Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules." Mol Biol Cell **13**(1): 195-210.
Kelly, M. P. (2017). A Role for Phosphodiesterase 11A (PDE11A) in the Formation of Social Memories and the Stabilization of Mood. Phosphodiesterases: CNS Functions and Diseases. H.-T. Zhang, Y. Xu and J. M. O'Donnell. Cham, Springer International Publishing**:** 201-230.
Kelly, M. P., S. F. Logue, J. Brennan, J. P. Day, S. Lakkaraju, L. Jiang, X. Zhong, M. Tam, S. J. Sukoff Rizzo, B. J. Platt, J. M. Dwyer, S. Neal, V. L. Pulito, M. J. Agostino, S. M. Grauer, R. L. Navarra, C. Kelley, T. A. Comery, R. J. Murrills, M. D. Houslay and N. J. Brandon (2010). "Phosphodiesterase 11A in brain is enriched in ventral hippocampus and deletion causes psychiatric disease-related phenotypes." Proc Natl Acad Sci U S A **107**(18): 8457-8462.
Kim, S. Y., J. Y. Kim, S. Malik, W. Son, K. S. Kwon and C. Kim (2012). "Negative regulation of EGFR/MAPK pathway by Pumilio in Drosophila melanogaster." PLoS One **7**(4): e34016.
Konarska, M. M., R. A. Padgett and P. A. Sharp (1984). "Recognition of cap structure in splicing in vitro of mRNA precursors." Cell **38**(3): 731-736.

Kondo, Y., C. Oubridge, A. M. van Roon and K. Nagai (2015). "Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition." Elife **4**.
Kopp, F., M. M. Elguindy, M. E. Yalvac, H. Zhang, B. Chen, F. A. Gillett, S. Lee, S. Sivakumar, H. Yu, Y. Xie, P. Mishra, Z. Sahenk and J. T. Mendell (2019). "PUMILIO hyperactivity drives premature aging of Norad-deficient mice." Elife **8**.
Kozak, M. (1986). "Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes." Cell **44**(2): 283-292.
Kraut, R., K. Menon and K. Zinn (2001). "A gain-of-function screen for genes controlling motor axon guidance and synaptogenesis in Drosophila." Current Biology **11**(6): 417-430.
Kuhn, U., M. Gundel, A. Knoth, Y. Kerwitz, S. Rudel and E. Wahle (2009). "Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor." J Biol Chem **284**(34): 22803-22814.
Lai, W. S., E. Carballo, J. M. Thorn, E. A. Kennington and P. J. Blackshear (2000). "Interactions of CCCH zinc finger proteins with mRNA. Binding of tristetraprolin-related zinc finger proteins to Au-rich elements and destabilization of mRNA." J Biol Chem **275**(23): 17827-17837.
Laughon, A., S. B. Carroll, F. A. Storfer, P. D. Riley and M. P. Scott (1985). "Common properties of proteins encoded by the Antennapedia complex genes of Drosophila melanogaster." Cold Spring Harb Symp Quant Biol **50**: 253-262.
Laver, J. D., X. Li, D. Ray, K. B. Cook, N. A. Hahn, S. Nabeel-Shah, M. Kekis, H. Luo, A. J. Marsolais, K. Y. Fung, T. R. Hughes, J. T. Westwood, S. S. Sidhu, Q. Morris, H. D. Lipshitz and C. A. Smibert (2015). "Brain tumor is a sequence-specific RNA-binding protein that directs maternal mRNA clearance during the Drosophila maternal-to-zygotic transition." Genome Biol **16**: 94.
Le Hir, H., E. Izaurralde, L. E. Maquat and M. J. Moore (2000). "The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions." EMBO J **19**(24): 6860-6869.
Lee, S., F. Kopp, T. C. Chang, A. Sataluri, B. Chen, S. Sivakumar, H. Yu, Y. Xie and J. T. Mendell (2016). "Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins." Cell **164**(1-2): 69-80.
Lee, S. H., C. Zhu, Y. Peng, D. T. Johnson, L. Lehmann and Z. Sun (2013). "Identification of a novel role of ZMIZ2 protein in regulating the activity of the Wnt/beta-catenin signaling pathway." J Biol Chem **288**(50): 35913-35924.
Leppek, K., R. Das and M. Barna (2018). "Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them." Nat Rev Mol Cell Biol **19**(3): 158-174.
Leppek, K., J. Schott, S. Reitter, F. Poetz, M. C. Hammond and G. Stoecklin (2013). "Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs." Cell **153**(4): 869-881.
Levy, I., E. Szarek, A. G. Maria, M. Starrost, M. De La Luz Sierra, F. R. Faucz and C. A. Stratakis (2021). "A phosphodiesterase 11 (Pde11a) knockout mouse expressed functional but reduced Pde11a: Phenotype and impact on adrenocortical function." Mol Cell Endocrinol **520**: 111071.

Lima, S. A., L. B. Chipman, A. L. Nicholson, Y. H. Chen, B. A. Yee, G. W. Yeo, J. Coller and A. E. Pasquinelli (2017). "Short poly(A) tails are a conserved feature of highly expressed genes." Nat Struct Mol Biol **24**(12): 1057-1063.
Lin, H. and A. C. Spradling (1997). "A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary." Development **124**(12): 2463-2476.
Lin, K., W. Qiang, M. Zhu, Y. Ding, Q. Shi, X. Chen, E. Zsiros, K. Wang, X. Yang, T. Kurita and E. Y. Xu (2019). "Mammalian Pum1 and Pum2 Control Body Size via Translational Regulation of the Cell Cycle Inhibitor Cdkn1b." Cell Rep **26**(9): 2434-2450 e2436.
Lin, K., S. Zhang, Q. Shi, M. Zhu, L. Gao, W. Xia, B. Geng, Z. Zheng and E. Y. Xu (2018). "Essential requirement of mammalian Pumilio family in embryonic development." Mol Biol Cell **29**(24): 2922-2932.
Lin, M. F., J. W. Carlson, M. A. Crosby, B. B. Matthews, C. Yu, S. Park, K. H. Wan, A. J. Schroeder, L. S. Gramates, S. E. St Pierre, M. Roark, K. L. Wiley, Jr., R. J. Kulathinal, P. Zhang, K. V. Myrick, J. V. Antone, S. E. Celniker, W. M. Gelbart and M. Kellis (2007). "Revisiting the protein-coding gene catalog of Drosophila melanogaster using 12 fly genomes." Genome Res **17**(12): 1823-1836.
Luo, Y., Z. Na and S. A. Slavoff (2018). "P-Bodies: Composition, Properties, and Functions." Biochemistry **57**(17): 2424-2431.
Lutz, C. S. (2008). "Alternative polyadenylation: a twist on mRNA 3' end formation." ACS Chem Biol **3**(10): 609-617.
Macdonald, P. M. (1992). "The Drosophila pumilio gene: an unusually long transcription unit and an unusual protein." Development **114**(1): 221-232.
Maillet, L., C. Tu, Y. K. Hong, E. O. Shuster and M. A. Collart (2000). "The essential function of Not1 lies within the Ccr4-Not complex." J Mol Biol **303**(2): 131-143.
Makhlouf, A., A. Kshirsagar and C. Niederberger (2006). "Phosphodiesterase 11: a brief review of structure, expression and function." Int J Impot Res **18**(6): 501-509.
Mandel, C. R., S. Kaneko, H. Zhang, D. Gebauer, V. Vethantham, J. L. Manley and L. Tong (2006). "Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease." Nature **444**(7121): 953-956.
Markmiller, S., S. Soltanieh, K. L. Server, R. Mak, W. Jin, M. Y. Fang, E. C. Luo, F. Krach, D. Yang, A. Sen, A. Fulzele, J. M. Wozniak, D. J. Gonzalez, M. W. Kankel, F. B. Gao, E. J. Bennett, E. Lecuyer and G. W. Yeo (2018). "Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules." Cell **172**(3): 590-604 e513.
Martinez, J. C., L. K. Randolph, D. M. Iascone, H. F. Pernice, F. Polleux and U. Hengst (2019). "Pum2 Shapes the Transcriptome in Developing Axons through Retention of Target mRNAs in the Cell Body." Neuron **104**(5): 931-946 e935.
Mathys, H., J. Basquin, S. Ozgur, M. Czarnocki-Cieciura, F. Bonneau, A. Aartse, A. Dziembowski, M. Nowotny, E. Conti and W. Filipowicz (2014). "Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression." Mol Cell **54**(5): 751-765.
Maurer, G., B. Tarkowski and M. Baccarini (2011). "Raf kinases in cancer-roles and therapeutic opportunities." Oncogene **30**(32): 3477-3488.
Mayr, C. (2017). "Regulation by 3'-Untranslated Regions." Annu Rev Genet **51**: 171-194.
Mayr, C. (2019). "What Are 3' UTRs Doing?" Cold Spring Harb Perspect Biol **11**(10).

Mayr, C. and D. P. Bartel (2009). "Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells." Cell **138**(4): 673-684.
McCracken, S., N. Fong, E. Rosonina, K. Yankulov, G. Brothers, D. Siderovski, A. Hessel, S. Foster, A. E. Program, S. Shuman and D. L. Bentley (1997). "5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II." Genes & Development **11**(24): 3306-3318.
Mee, C. J., E. C. Pym, K. G. Moffat and R. A. Baines (2004). "Regulation of neuronal excitability through pumilio-dependent control of a sodium channel gene." J Neurosci **24**(40): 8695-8703.
Menon, K. P., S. Andrews, M. Murthy, E. R. Gavis and K. Zinn (2009). "The translational repressors Nanos and Pumilio have divergent effects on presynaptic terminal growth and postsynaptic glutamate receptor subunit composition." J Neurosci **29**(17): 5558-5572.
Menon, K. P., S. Sanyal, Y. Habara, R. Sanchez, R. P. Wharton, M. Ramaswami and K. Zinn (2004). "The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E." Neuron **44**(4): 663-676.
Michellod, M. A. and N. B. Randsholt (2008). "Implication of the Drosophila beta-amyloid peptide binding-like protein AMX in Notch signaling during early neurogenesis." Brain Res Bull **75**(2-4): 305-309.
Miles, W. O., K. Tschop, A. Herr, J. Y. Ji and N. J. Dyson (2012). "Pumilio facilitates miRNA regulation of the E2F3 oncogene." Genes Dev **26**(4): 356-368.
Mittag, T. and R. Parker (2018). "Multiple Modes of Protein-Protein Interactions Promote RNP Granule Assembly." J Mol Biol **430**(23): 4636-4649.
Monribot-Villanueva, J., R. A. Juarez-Uribe, Z. Palomera-Sanchez, L. Gutierrez-Aguiar, M. Zurita, J. A. Kennison and M. Vazquez (2013). "TnaA, an SP-RING protein, interacts with Osa, a subunit of the chromatin remodeling complex BRAHMA and with the SUMOylation pathway in Drosophila melanogaster." PLoS One **8**(4): e62251.
Morris, A. R., N. Mukherjee and J. D. Keene (2008). "Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets." Mol Cell Biol **28**(12): 4093-4103.
Muraro, N. I., A. J. Weston, A. P. Gerber, S. Luschnig, K. G. Moffat and R. A. Baines (2008). "Pumilio binds para mRNA and requires Nanos and Brat to regulate sodium current in Drosophila motoneurons." J Neurosci **28**(9): 2099-2109.
Murata, Y. and R. P. Wharton (1995). "Binding of pumilio to maternal hunchback mrna is required for posterior patterning in drosophila embryos." Cell **80**: 747-756.
Narita, R., K. Takahasi, E. Murakami, E. Hirano, S. P. Yamamoto, M. Yoneyama, H. Kato and T. Fujita (2014). "A novel function of human Pumilio proteins in cytoplasmic sensing of viral infection." PLoS Pathog **10**(10): e1004417.
Nasertorabi, F., C. Batisse, M. Diepholz, D. Suck and B. Bottcher (2011). "Insights into the structure of the CCR4-NOT complex by electron microscopy." FEBS Lett **585**(14): 2182-2186.
Naudin, C., A. Hattabi, F. Michelet, A. Miri-Nezhad, A. Benyoucef, F. Pflumio, F. Guillonneau, S. Fichelson, I. Vigon, I. Dusanter-Fourt and E. Lauret (2017). "PUMILIO/FOXP1 signaling drives expansion of hematopoietic stem/progenitor and leukemia cells." Blood **129**(18): 2493-2506.
Nicholson, A. L. and A. E. Pasquinelli (2019). "Tales of Detailed Poly(A) Tails." Trends Cell Biol **29**(3): 191-200.

Nishida, Y., M. Hata, T. Ayaki, H. Ryo, M. Yamagata, K. Shimizu and Y. Nishizuka (1988). "Proliferation of both somatic and germ cells is affected in the Drosophila mutants of raf proto-oncogene." The EMBO Journal **7**(3): 775-781.
Nusslein-Volhard, C., H. G. Frohnhofer and R. Lehmann (1987). "Determination of anteroposterior polarity in Drosophila." Science **238**(4834): 1675-1681.
Oh, S., S. Shin and R. Janknecht (2012). "ETV1, 4 and 5: an oncogenic subfamily of ETS transcription factors." Biochim Biophys Acta **1826**(1): 1-12.
Ohashi, R. and N. Shiina (2020). "Cataloguing and Selection of mRNAs Localized to Dendrites in Neurons and Regulated by RNA-Binding Proteins in RNA Granules." Biomolecules **10**(2).
Ohno, M., H. Sakamoto and Y. Shimura (1987). "Preferential excision of the 5' proximal intron from mRNA precursors with two introns as mediated by the cap structure." Proc Natl Acad Sci U S A **84**(15): 5187-5191.
Olivas, W. and R. Parker (2000). "The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast." EMBO J **19**(23): 6602-6611.
Owusu-Ansah, E. and U. Banerjee (2009). "Reactive oxygen species prime Drosophila haematopoietic progenitors for differentiation." Nature **461**(7263): 537-541.
Ozgur, S., J. Basquin, A. Kamenska, W. Filipowicz, N. Standart and E. Conti (2015). "Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex." Cell Rep **13**(4): 703-711.
Parisi, M. and H. F. Lin (1999). "The Drosophila pumilio gene encodes two functional protein isoforms that play multiple roles in germline development, gonadogenesis, oogenesis and embryogenesis." Genetics **153**(1): 235-250.
Paul, S. M., M. J. Palladino and G. J. Beitel (2007). "A pump-independent function of the Na,K-ATPase is required for epithelial junction function and tracheal tube-size control." Development **134**(1): 147-155.
Peng, H., L. Wang, Q. Su, K. Yi, J. Du and Z. Wang (2019). "MiR-31-5p promotes the cell growth, migration and invasion of colorectal cancer cells by targeting NUMB." Biomed Pharmacother **109**: 208-216.
Petit, A. P., L. Wohlbold, P. Bawankar, E. Huntzinger, S. Schmidt, E. Izaurralde and O. Weichenrieder (2012). "The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex." Nucleic Acids Res **40**(21): 11058-11072.
Pillutla, R. C., A. Shimamoto, Y. Furuichi and A. J. Shatkin (1998). "Human mRNA capping enzyme (RNGTT) and cap methyltransferase (RNMT) map to 6q16 and 18p11.22-p11.23, respectively." Genomics **54**(2): 351-353.
Pinnell, N., R. Yan, H. J. Cho, T. Keeley, M. J. Murai, Y. Liu, A. S. Alarcon, J. Qin, Q. Wang, R. Kuick, K. S. Elenitoba-Johnson, I. Maillard, L. C. Samuelson, T. Cierpicki and M. Y. Chiang (2015). "The PIAS-like Coactivator Zmiz1 Is a Direct and Selective Cofactor of Notch1 in T Cell Development and Leukemia." Immunity **43**(5): 870-883.
Pique, M., J. M. Lopez, S. Foissac, R. Guigo and R. Mendez (2008). "A combinatorial code for CPE-mediated translational control." Cell **132**(3): 434-448.
Raisch, T., D. Bhandari, K. Sabath, S. Helms, E. Valkov, O. Weichenrieder and E. Izaurralde (2016). "Distinct modes of recruitment of the CCR4-NOT complex by Drosophila and vertebrate Nanos." EMBO J **35**(9): 974-990.
Raisch, T., C. T. Chang, Y. Levdansky, S. Muthukumar, S. Raunser and E. Valkov (2019). "Reconstitution of recombinant human CCR4-NOT reveals molecular insights into regulated deadenylation." Nat Commun **10**(1): 3173.

Ramanathan, A., G. B. Robb and S. H. Chan (2016). "mRNA capping: biological functions and applications." Nucleic Acids Res **44**(16): 7511-7526.
Richter, J. D. and P. Lasko (2011). "Translational control in oocyte development." Cold Spring Harb Perspect Biol **3**(9): a002758.
Rodrigues, J. P., M. Rode, D. Gatfield, B. J. Blencowe, M. Carmo-Fonseca and E. Izaurralde (2001). "REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus." Proc Natl Acad Sci U S A **98**(3): 1030-1035.
Rosales-Vega, M., A. Hernandez-Becerril, J. M. Murillo-Maldonado, M. Zurita and M. Vazquez (2018). "The role of the trithorax group TnaA isoforms in Hox gene expression, and in Drosophila late development." PLoS One **13**(10): e0206587.
Routh, A. (2019). "DPAC: A Tool for Differential Poly(A)-Cluster Usage from Poly(A)-Targeted RNAseq Data." G3 (Bethesda) **9**(6): 1825-1830.
Routh, A., S. R. Head, P. Ordoukhanian and J. E. Johnson (2015). "ClickSeq: Fragmentation-Free Next-Generation Sequencing via Click Ligation of Adaptors to Stochastically Terminated 3'-Azido cDNAs." J Mol Biol **427**(16): 2610-2616.
Routh, A., P. Ji, E. Jaworski, Z. Xia, W. Li and E. J. Wagner (2017). "Poly(A)-ClickSeq: click-chemistry for next-generation 3-end sequencing without RNA enrichment or fragmentation." Nucleic Acids Res **45**(12): e112.
Ryan, K., O. Calvo and J. L. Manley (2004). "Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease." RNA **10**(4): 565-573.
Salazar, A. M., E. J. Silverman, K. P. Menon and K. Zinn (2010). "Regulation of synaptic Pumilio function by an aggregation-prone domain." J Neurosci **30**(2): 515-522.
Sandler, H., J. Kreth, H. T. Timmers and G. Stoecklin (2011). "Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin." Nucleic Acids Res **39**(10): 4373-4386.
Sanduja, S., F. F. Blanco, L. E. Young, V. Kaza and D. A. Dixon (2012). "The role of tristetraprolin in cancer and inflammation." Front Biosci (Landmark Ed) **17**: 174-188.
Sato, H. and L. E. Maquat (2009). "Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta." Genes Dev **23**(21): 2537-2550.
Sato, T., Russell, M., Denell, R. E.. (1983). "Homoeosis in Drosophila: A new enhancer of Polycomb and related Homoeotic mutations." Genetics **105**(2): 357-370.
Schwanhausser, B., D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen and M. Selbach (2011). "Global quantification of mammalian gene expression control." Nature **473**(7347): 337-342.
Schweers, B. A., K. J. Walters and M. Stern (2002). "The Drosophila melanogaster translational repressor pumilio regulates neuronal excitability." Genetics **161**(3): 1177-1185.
Searle, N. E. and L. Pillus (2018). "Critical genomic regulation mediated by Enhancer of Polycomb." Curr Genet **64**(1): 147-154.
Sgromo, A., T. Raisch, C. Backhaus, C. Keskeny, V. Alva, O. Weichenrieder and E. Izaurralde (2018). "Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets." RNA **24**(3): 381-395.
Sgromo, A., T. Raisch, P. Bawankar, D. Bhandari, Y. Chen, D. Kuzuoglu-Ozturk, O. Weichenrieder and E. Izaurralde (2017). "A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin." Nat Commun **8**: 14307.

Sharif, H., S. Ozgur, K. Sharma, C. Basquin, H. Urlaub and E. Conti (2013). "Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions." Nucleic Acids Res **41**(17): 8377-8390.
Shi, Y., B. E. Paluch, X. Wang and X. Jiang (2012). "PTEN at a glance." J Cell Sci **125**(Pt 20): 4687-4692.
Shimberg, G. D., J. L. Michalek, A. A. Oluyadi, A. V. Rodrigues, B. E. Zucconi, H. M. Neu, S. Ghosh, K. Sureschandra, G. M. Wilson, T. L. Stemmler and S. L. Michel (2016). "Cleavage and polyadenylation specificity factor 30: An RNA-binding zinc-finger protein with an unexpected 2Fe-2S cluster." Proc Natl Acad Sci U S A **113**(17): 4700-4705.
Shuai, K. and B. Liu (2005). "Regulation of gene-activation pathways by PIAS proteins in the immune system." Nat Rev Immunol **5**(8): 593-605.
Simms, C. L., E. N. Thomas and H. S. Zaher (2017). "Ribosome-based quality control of mRNA and nascent peptides." Wiley Interdiscip Rev RNA **8**(1).
Smith, B. L., D. R. Gallie, H. Le and P. K. Hansma (1997). "Visualization of poly(A)-binding protein complex formation with poly(A) RNA using atomic force microscopy." J Struct Biol **119**(2): 109-117.
Sonenberg, N. and A. G. Hinnebusch (2009). "Regulation of translation initiation in eukaryotes: mechanisms and biological targets." Cell **136**(4): 731-745.
Stepien, B. K., C. Oppitz, D. Gerlach, U. Dag, M. Novatchkova, S. Kruttner, A. Stark and K. Keleman (2016). "RNA-binding profiles of Drosophila CPEB proteins Orb and Orb2." Proc Natl Acad Sci U S A **113**(45): E7030-E7038.
Stewart, G. L., K. S. S. Enfield, A. P. Sage, V. D. Martinez, B. C. Minatel, M. E. Pewarchuk, E. A. Marshall and W. L. Lam (2019). "Aberrant Expression of Pseudogene-Derived lncRNAs as an Alternative Mechanism of Cancer Gene Regulation in Lung Adenocarcinoma." Front Genet **10**: 138.
Subtelny, A. O., S. W. Eichhorn, G. R. Chen, H. Sive and D. P. Bartel (2014). "Poly(A)-tail profiling reveals an embryonic switch in translational control." Nature **508**(7494): 66-71.
Sun, X., G. Zhangyuan, L. Shi, Y. Wang, B. Sun and Q. Ding (2017). "Prognostic and clinicopathological significance of cyclin B expression in patients with breast cancer: A meta-analysis." Medicine (Baltimore) **96**(19): e6860.
Szarek, E. and C. A. Stratakis (2014). "Phosphodiesterases and adrenal Cushing in mice and humans." Horm Metab Res **46**(12): 863-868.
Tadros, W. and H. D. Lipshitz (2009). "The maternal-to-zygotic transition: a play in two acts." Development **136**(18): 3033-3042.
Tan, D., M. Zhou, M. Kiledjian and L. Tong (2014). "The ROQ domain of Roquin recognizes mRNA constitutive-decay element and double-stranded RNA." Nat Struct Mol Biol **21**(8): 679-685.
Tatsumi, K., S. Tsuji, H. Miwa, T. Morisaku, M. Nuriya, M. Orihara, K. Kaneko, H. Okano and M. Yasui (2009). "Drosophila big brain does not act as a water channel, but mediates cell adhesion." FEBS Lett **583**(12): 2077-2082.
Temme, C., S. Zaessinger, S. Meyer, M. Simonelig and E. Wahle (2004). "A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in Drosophila." EMBO J **23**(14): 2862-2871.
Temme, C., L. Zhang, E. Kremmer, C. Ihling, A. Chartier, A. Sinz, M. Simonelig and E. Wahle (2010). "Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation." RNA **16**(7): 1356-1370.

Texada, M. J., T. Koyama and K. Rewitz (2020). "Regulation of Body Size and Growth Control." Genetics **216**(2): 269-313.
Thurmond J, G. J., Strelets VB, Attrill H, Gramates LS, Marygold SJ, Matthews BB, Millburn G, Antonazzo G, Trovisco V, Kaufman TC, Calvi BR and the FlyBase Consortium (2019). "FlyBase 2.0: the next generation." Nucleic Acids Res **47**(D1): D759-D765.
Tichon, A., N. Gil, Y. Lubelsky, T. Havkin Solomon, D. Lemze, S. Itzkovitz, N. Stern-Ginossar and I. Ulitsky (2016). "A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells." Nat Commun **7**: 12209.
Tosoni, D., S. Zecchini, M. Coazzoli, I. Colaluca, G. Mazzarol, A. Rubio, M. Caccia, E. Villa, O. Zilian, P. P. Di Fiore and S. Pece (2015). "The Numb/p53 circuitry couples replicative self-renewal and tumor suppression in mammary epithelial cells." J Cell Biol **211**(4): 845-862.
Uemura, T., S. Shepherd, L. Ackerman, L. Y. Jan and Y. N. Jan (1989). "numb, a gene required in determination of cell fate during sensory organ formation in Drosophila embryos." Cell **58**(2): 349-360.
Ukleja, M., J. Cuellar, A. Siwaszek, J. M. Kasprzak, M. Czarnocki-Cieciura, J. M. Bujnicki, A. Dziembowski and J. M. Valpuesta (2016). "The architecture of the Schizosaccharomyces pombe CCR4-NOT complex." Nat Commun **7**: 10433.
Uversky, V. N. (2019). "Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics." Frontiers in Physics **7**.
van der Lee, R., M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright and M. M. Babu (2014). "Classification of intrinsically disordered regions and proteins." Chem Rev **114**(13): 6589-6631.
Van Etten, J., T. L. Schagat, J. Hrit, C. A. Weidmann, J. Brumbaugh, J. J. Coon and A. C. Goldstrohm (2012). "Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs." J Biol Chem **287**(43): 36370-36383.
Vardy, L. and T. L. Orr-Weaver (2007). "The Drosophila PNG kinase complex regulates the translation of cyclin B." Dev Cell **12**(1): 157-166.
Vessey, J. P., A. Vaccani, Y. Xie, R. Dahm, D. Karra, M. A. Kiebler and P. Macchi (2006). "Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules." J Neurosci **26**(24): 6496-6508.
Vicens, Q., J. S. Kieft and O. S. Rissland (2018). "Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication." Mol Cell **72**(5): 805-812.
Wahle, E. (1995). "Poly(A) tail length control is caused by termination of processive synthesis." Journal of Biological Chemistry **270**(6): 2800-2808.
Wahle, E. and G. S. Winkler (2013). "RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes." Biochim Biophys Acta **1829**(6-7): 561-570.
Walter, P. and G. Blobel (1981). "Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes." J Cell Biol **91**(2 Pt 1): 557-561.
Wang, H., D. Xiang, B. Liu, A. He, H. J. Randle, K. X. Zhang, A. Dongre, N. Sachs, A. P. Clark, L. Tao, Q. Chen, V. V. Botchkarev, Jr., Y. Xie, N. Dai, H. Clevers, Z. Li and D. M. Livingston (2019). "Inadequate DNA Damage Repair Promotes Mammary Transdifferentiation, Leading to BRCA1 Breast Cancer." Cell **178**(1): 135-151 e119.

Wang, X., J. Y. Yang, J. Cai, D. J. Zhang, L. Zhao, L. H. Luo, Y. Xiong, T. Zhang and M. Jin (2021). "MiR-543/Numb promotes proliferation, metastasis, and stem-like cell traits of prostate cancer cells." Am J Transl Res **13**(2): 617-631.
Wang, X., P. D. Zamore and T. M. Hall (2001). "Crystal structure of a Pumilio homolgy domain." Molecular Cell **7**: 855-865.
Wayman, C., S. Phillips, C. Lunny, T. Webb, L. Fawcett, R. Baxendale and G. Burgess (2005). "Phosphodiesterase 11 (PDE11) regulation of spermatozoa physiology." Int J Impot Res **17**(3): 216-223.
Weidmann, C. A. and A. C. Goldstrohm (2012). "Drosophila Pumilio protein contains multiple autonomous repression domains that regulate mRNAs independently of Nanos and brain tumor." Mol Cell Biol **32**(2): 527-540.
Weidmann, C. A., C. Qiu, R. M. Arvola, T. F. Lou, J. Killingsworth, Z. T. Campbell, T. M. Tanaka Hall and A. C. Goldstrohm (2016). "Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio." Elife **5**.
Weidmann, C. A., N. A. Raynard, N. H. Blewett, J. Van Etten and A. C. Goldstrohm (2014). "The RNA binding domain of Pumilio antagonizes poly-adenosine binding protein and accelerates deadenylation." RNA **20**(8): 1298-1319.
Weston, A. J. and R. A. Baines (2007). "Translational regulation of neuronal electrical properties." Invert Neurosci **7**(2): 75-86.
Weyn-Vanhentenryck, S. M., A. Mele, Q. Yan, S. Sun, N. Farny, Z. Zhang, C. Xue, M. Herre, P. A. Silver, M. Q. Zhang, A. R. Krainer, R. B. Darnell and C. Zhang (2014). "HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism." Cell Rep **6**(6): 1139-1152.
Wharton, R. P., J. Sonoda, T. Lee, M. Patterson and Y. Murata (1998). "The Pumilio RNA-binding domain is also a translational regulator." Molecular Cell **1**(6): 863-872.
Wickens, M., D. S. Bernstein, J. Kimble and R. Parker (2002). "A PUF family portrait: 3′UTR regulation as a way of life." Trends in Genetics **18**(3): 150-157.
Willis, I. M. (1993). "RNA polymerase III. Genes, factors and transcriptional specificity." Eur J Biochem **212**(1): 1-11.
Winkler, G. S. (2010). "The mammalian anti-proliferative BTG/Tob protein family." J Cell Physiol **222**(1): 66-72.
Wolfe, M. B., T. L. Schagat, M. T. Paulsen, B. Magnuson, M. Ljungman, D. Park, C. Zhang, Z. T. Campbell, A. C. Goldstrohm and P. L. Freddolino (2020). "Principles of mRNA control by human PUM proteins elucidated from multi-modal experiments and integrative data analysis." RNA.
Wolozin, B. and P. Ivanov (2019). "Stress granules and neurodegeneration." Nat Rev Neurosci **20**(11): 649-666.
Wreden, C., A. C. Verrotti, J. A. Schisa, M. E. Lieberfarb and S. Strickland (1997). "Nanos and pumilio establish embryonic polarity in Drosophila by promoting posterior deadenylation of hunchback mRNA." Development **124**(15): 3015-3023.
Xing, W., D. Muhlrad, R. Parker and M. K. Rosen (2020). "A quantitative inventory of yeast P body proteins reveals principles of composition and specificity." Elife **9**.
Xu, P., B. Sun and P. M. Salvaterra (1999). "Organization and transcriptional regulation of Drosophila Na+,K+-ATPase β subunit genes: Nrv1 and Nrv2." Gene **236**(2): 303-313.
Yamada-Okabe, T., R. Doi, O. Shimmi, M. Arisawa and H. Yamada-Okabe (1998). "Isolation and characterization of a human cDNA for mRNA 5'-capping enzyme." Nucleic Acids Res **26**(7): 1700-1706.

Yang, G., R. Zhou, Q. Zhou, X. Guo, C. Yan, M. Ke, J. Lei and Y. Shi (2019). "Structural basis of Notch recognition by human gamma-secretase." Nature **565**(7738): 192-197.
Ye, B., C. Petritsch, I. E. Clark, E. R. Gavis, L. Y. Jan and Y. N. Jan (2004). "Nanos and Pumilio are essential for dendrite morphogenesis in Drosophila peripheral neurons." Curr Biol **14**(4): 314-321.
Yi, H., J. Park, M. Ha, J. Lim, H. Chang and V. N. Kim (2018). "PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay." Mol Cell **70**(6): 1081-1088 e1085.
Young, R. A. (1991). "RNA polymerase II." Annu Rev Biochem **60**: 689-715.
Zamore, P. D., J. R. Williamson and R. Lehmann (1997). "The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins." Rna-a Publication of the Rna Society **3**(12): 1421-1433.
Zarkower, D. and J. Hodgkin (1992). "Molecular analysis of the C. elegans sex-determining gene tra-1: A gene encoding two zinc finger proteins." Cell **70**(2): 237-249.
Zhai, B., J. Villen, S. A. Beausoleil, J. Mintseris and S. P. Gygi (2008). "Phosphoproteome analysis of Drosophila melanogaster embryos." J Proteome Res **7**(4): 1675-1682.
Zhang, B. L., M. Gallegos, A. Puoti, E. Durkin, S. Fields, J. Kimble and M. P. Wickens (1997). "A conserved RNA-binding protein that regulates sexual fates in the C-elegans hermaphrodite germ line." Nature **390**(6659): 477-484.
Zhang, M., D. Chen, J. Xia, W. Han, X. Cui, N. Neuenkirchen, G. Hermes, N. Sestan and H. Lin (2017). "Post-transcriptional regulation of mouse neurogenesis by Pumilio proteins." Genes Dev.
Zhang, X., Q. Li, L. Wang, Z. J. Liu and Y. Zhong (2018). "Active Protection: Learning-Activated Raf/MAPK Activity Protects Labile Memory from Rac1-Independent Forgetting." Neuron **98**(1): 142-155 e144.
Zhang, X., Y. Li, H. Xu and Y. W. Zhang (2014). "The gamma-secretase complex: from structure to function." Front Cell Neurosci **8**: 427.
Zhao, Y., T. Kiss, J. DelFavero, L. Li, X. Li, L. Zheng, J. Wang, C. Jiang, J. Shi, Z. Ungvari, A. Csiszar and X. A. Zhang (2020). "CD82-TRPM7-Numb signaling mediates age-related cognitive impairment." Geroscience **42**(2): 595-611.

Appendices

Appendix A: ConSurf analysis of RD3

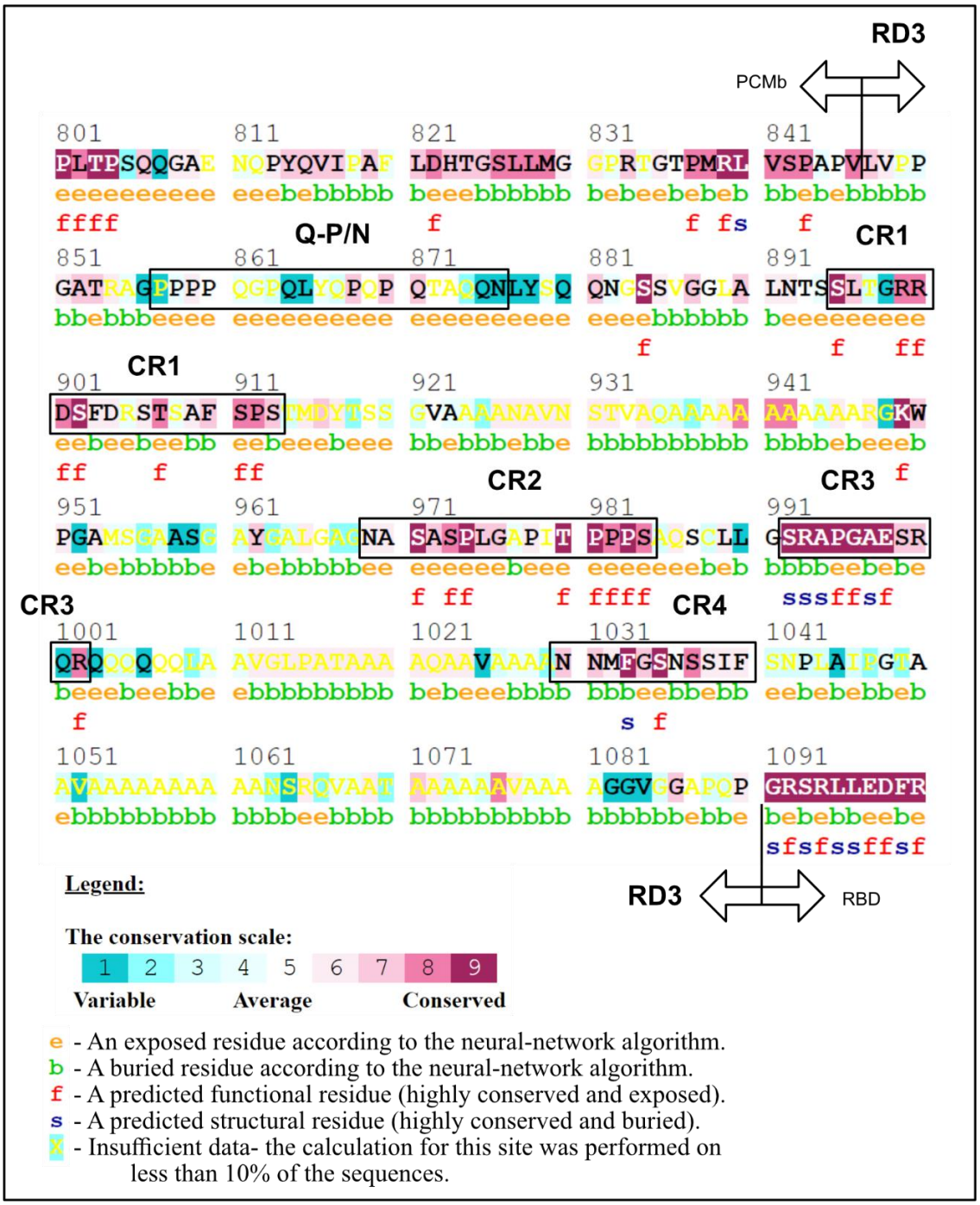


Figure A1: ConSurf conservation analysis of RD3. Analysis generated from an alignment of Pum proteins from 45 different species. CRs in RD3 are boxed. PCMb and the RBD are indicated. Base image taken from HTML document generated by ConSurf.

Appendix B: Antibodies used

Antibody	Supplier	Catalog number	Dilution used	Diluent
Mouse anti-V5	Invitrogen	R960-25	1:5000	Blotto
Rabbit anti-V5	CST	13202S	1:5000	TBST
Mouse anti-Flag	Sigma	F3165-1MG	1:5000	Blotto
Rabbit anti-Flag	Sigma	sab4301135	1:5000	Blotto
Rabbit anti-HA	CST	3724s	1:5000	TBST
Rabbit anti-Myc	CST	2278S	1:5000	TBST
Mouse anti-Tubulin	CST	3873	1:1000	Blotto
Rabbit anti-Vinculin	Thermo Fisher	700062	1:1000	Blotto
Goat anti-mouse HRP	Thermo Fisher	31430	1:5000	Blotto
Goat anti-rabbit HRP	CST	7074P2	1:5000	Blotto
Goat anti-rabbit HRP	Sigma	AP187P	1:5000	Blotto

Table B1: Antibodies used in this study. Composition of Blotto: 1x PBS, 0.1% Tween, 5% dried milk. TBST: 1x TBS, 5% BSA, 0.1% Tween.

Appendix C: Plasmids

Archive number	Plasmid name	How cloned	Creator	Primers
CAW061	pIZ MS2CP-GFP V5H6	restriction cloned	Chase Weidmann	CW115/116 for insert
RMA053	pIZ MS2-Dcp1 V5H6	restriction cloned from cDNA	Rene Arvola	
CAW314	pIZ MS2 Pum RD3i V5H6	inverse PCR	Chase Weidmann	
CAW313	pIZ MS2 Pum RD3ii V5H6	inverse PCR	Chase Weidmann	
RMA008	pIZ Dm Pum RD3-RBD	inverse PCR	Rene Arvola	R: CW 055 F: CW 244
pBA017	pAc5.4 NLuc 2xMS2 BS	Oligo cloning using AG 784-785	Brian Alzua	
CAW023	pAc5.1 FFLuc2 minimal 3'-UTR	oligo cloned	Chase Weidmann	
RJH63	pIZ MS2 RD3 ▲FA2	InvPCR from RD3	Rebecca Haugen	RJH 100+101

Archive number	Plasmid name	How cloned	Creator	Primers
RJH64	pIZ MS2 RD3 ▲Q+FA2	invPCR from RD3 ▲QN	Rebecca Haugen	RJH 100+101
RMA001	pIZ MS2-RD3 -QN	INV PCR/PNK/T4	Rene Arvola	R: RA 001 F: RA 002
RJH66	pIZ MS2 RD3 ▲FA1+CR3	invPCR from RD3 ▲CR3	Rebecca Haugen	RJH 98+99
RJH67	pIZ MS2 RD3 ▲FA2+CR3	invPCR from RD3 ▲CR3	Rebecca Haugen	RJH 100+101
RJH68	pIZ MS2 RD3 ▲FA4+CR3	invPCR from RD3 ▲CR3	Rebecca Haugen	RJH 102+103
RJH72	pIZ MS2 RD3 ▲1/3	invPCR from RD3	Rebecca Haugen	RJH 107+108
RJH73	pIZ MS2 RD3 ▲2/3	invPCR from RD3	Rebecca Haugen	RJH 106+105
RJH77	pIZ MS2 RD3 ▲FA1	invPCR from RD3	Rebecca Haugen	RJH 98+99
RJH78	pIZ MS2 RD3 ▲FA4	invPCR from RD3	Rebecca Haugen	RJH 102+103
RJH80	pIZ MS2 RD3 1/3 alone, 1-79	invPCR from RD3	Rebecca Haugen	RJH 104+105
RJH81	pIZ MS2 RD3 3/3 alone, 163-243	invPCR from RD3	Rebecca Haugen	RJH 106+107
RJH82	pIZ MS2 RD3 ▲3/3	invPCR from RD3	Rebecca Haugen	RJH 109+104
RJH87	pIZ MS2 RD3 ▲FA2+1	InvPCR from RD3 ▲FA2	Rebecca Haugen	RJH 98+99
RJH88	pIZ MS2 RD3 ▲FA2+4	InvPCR from RD3 ▲FA2	Rebecca Haugen	RJH 102+103
RJH98	pIZ MS2 RD3 2/3 alone, 80-162	invPCR from RD3 ▲1/3	Rebecca Haugen	RJH 104+109
RJH101	pIZ MS2 RD3 ▲FA1+4	invPCR from RD3 ▲FA4	Rebecca Haugen	RJH 98+99
RJH115	pIZ MS2 RD3 1-195	invPCR from RD3	Rebecca Haugen	RJH 113+104
RJH116	pIZ MS2 RD3 73-195	invPCR from RD3 73-243	Rebecca Haugen	RJH 113+104
RJH135	pIZ MS2 RD3 CR4 NN184-185AA	QuikChange from full-length RD3	Rebecca Haugen	RJH 116+117
RJH136	pIZ MS2 RD3 CR4 MF186-187AA	QuikChange from full-length RD3	Rebecca Haugen	RJH 118+119
RJH137	pIZ MS2 RD3 CR4 GS188-189AA	QuikChange from full-length RD3	Rebecca Haugen	RJH 120+121
RJH138	pIZ MS2 RD3 CR4 NS190-191AA	QuikChange from full-length RD3	Rebecca Haugen	RJH 122+123

Archive number	Plasmid name	How cloned	Creator	Primers
RJH139	pIZ MS2 RD3 CR4 SIF192-194AAA	QuikChange from full-length RD3	Rebecca Haugen	RJH 124+125
RJH146	pIZ MS2 RD3 CR4 F187A	QuikChange from full-length RD3	Rebecca Haugen	RJH 142+143
RJH147	pIZ MS2 RD3 CR4 F194A	QuikChange from full-length RD3	Rebecca Haugen	RJH 144+145
RJH148	pIZ MS2 RD3 CR1 SLTG-A's	QuikChange from full-length RD3	Rebecca Haugen	RJH 146+147
RJH149	pIZ MS2 RD3 CR1 RRD-A's	QuikChange from full-length RD3	Rebecca Haugen	RJH 148+149
RJH150	pIZ MS2 RD3 CR1 SFD-A's	QuikChange from full-length RD3	Rebecca Haugen	RJH 150+151
RJH151	pIZ MS2 RD3 CR1 RST-A's	QuikChange from full-length RD3	Rebecca Haugen	RJH 152+153
RJH152	pIZ MS2 RD3 CR1 SAF-A's	QuikChange from full-length RD3	Rebecca Haugen	RJH 154+155
RJH153	pIZ MS2 RD3 CR1 SPS-A's	QuikChange from full-length RD3	Rebecca Haugen	RJH 156+157
RJH158	pIZ MS2 eGFP MCS	Inverse PCR to add BglII and PvuI restriction sites	Rebecca Haugen	RJH 166+168
RJH161	pIZ MS2 RD3 CR4 F187+194A	QuikChange from RD3 F187A to mutate F194	Rebecca Haugen	RJH 144+145
RJH162	pIZ MS2 RD3 CR4 SSS-A's	QuikChange from full-length RD3	Rebecca Haugen	RJH 188+189
RJH163	pIZ MS2 RD3 CR1 SLTG+CR4 F187A	QuikChange from RD3 CR1 SLTG-A's	Rebecca Haugen	RJH 142+143
RJH164	pIZ MS2 RD3 CR1 RRD+CR4 F187A	QuikChange from RD3 CR1 RRD-A's	Rebecca Haugen	RJH 142+143
RJH165	pIZ MS2 RD3 CR1 SFD+CR4 F187A	QuikChange from RD3 CR1 SFD-A's	Rebecca Haugen	RJH 142+143
RJH166	pIZ MS2 RD3 CR1 RST+CR4 F187A	QuikChange from RD3 CR1 RST-A's	Rebecca Haugen	RJH 142+143
RJH171	pIZ MS2 RD3 49-195	Inverse PCR from RD3 1-195	Rebecca Haugen	RJH 107+190
RJH172	pIZ MS2 RD3 CR1 SPS+CR4 F187A	QuikChange from RD3 CR1 SPS-A's	Rebecca Haugen	RJH 142+143
RJH173	pIZ MS2 RD3 CR1 SAF+CR4 F187A	QuikChange from RD3 CR1 SAF-A's	Rebecca Haugen	RJH 142+143
RJH240	pIZ MS2-eGFP-link-CR1-link-CR4 (1x)	gBlock cloned	Rebecca Haugen	2-linker CR1+4 gBlock

Archive number	Plasmid name	How cloned	Creator	Primers
RJH241	pIZ MS2-eGFP-link-CR1-link-CR4 (2x)	gBlock cloned w/ restriction sites to multimerize	Rebecca Haugen	2-linker CR1+4 gBlock
RJH242	pIZ MS2-eGFP-link-CR1-link-CR4 (3x)	gBlock cloned w/ restriction sites to multimerize	Rebecca Haugen	2-linker CR1+4 gBlock
RJH243	pIZ MS2-eGFP-link-CR1-link-CR4 (4x)	gBlock cloned w/ restriction sites to multimerize	Rebecca Haugen	2-linker CR1+4 gBlock
RJH184	pIZ MS2-eGFP-link-CR1-link-CR4 (5x)	gBlock cloned w/ restriction sites to multimerize	Rebecca Haugen	2-linker CR1+4 gBlock
RJH185	pIZ MS2-eGFP-link-CR1-link-CR4 (6x)	gBlock cloned w/ restriction sites to multimerize	Rebecca Haugen	2-linker CR1+4 gBlock
RJH186	pIZ MS2-eGFP-link-CR1-link-CR4 (5x) FF mut	gBlock cloned w/ restriction sites to multimerize	Rebecca Haugen	2-linker CR1+4 gBlock
RJH187	pIZ MS2-eGFP-RD3	PCR/restriction cloned	Rebecca Haugen	RJH 199+200
RJH188	pGBK RD3	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH189	pGBK RD3 ΔQ	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH190	pGBK RD3 ΔCR1	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH191	pGBK RD3 ΔCR2	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH192	pGBK RD3 ΔCR3	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH193	pGBK RD3 ΔCR4	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH194	pGBK RD3 ΔCR1+4	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH195	pGBK RD3 F1033A	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH196	pGBK RD3 F1040A	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
RJH197	pGBK RD3 FF	PCR/restriction cloned	Rebecca Haugen	RJH 312+313
CAW038	pIZ MS2CP-DmPop2 V5H6	Amplified from S2 cDNA w/ CW095, 096 cut with BamHI and XbaI, ligated CAW001	Chase Weidmann	AG871/872

Archive number	Plasmid name	How cloned	Creator	Primers
ACG272	pGADT7		Clonetech	
ACG1036	pGADT7 T		CloneTech	
ACG1037	pGBKT7 p53		CloneTech	
RJH198	pGAD Not1-N	PCR/restriction cloned	Rebecca Haugen	RJH 337+338
RJH199	pGAD Not1-C	PCR/restriction cloned	Rebecca Haugen	RJH 335+336
RJH200	pGAD Not1-M	PCR/restriction cloned	Rebecca Haugen	RJH 339+340
RJH201	pGAD Dcp1	PCR/restriction cloned	Rebecca Haugen	RJH 383+384
RJH202	pGAD Me31b	PCR/restriction cloned	Rebecca Haugen	RJH 381+382
RJH203	pGAD Tis11	PCR/restriction cloned	Rebecca Haugen	RJH 399+400
RJH204	pGAD CCR4	PCR/restriction cloned	Rebecca Haugen	RJH 397+398
RJH205	pGAD Pop2	PCR/restriction cloned	Rebecca Haugen	RJH 341+342
RJH206	pGAD Not10	PCR/restriction cloned	Rebecca Haugen	RJH 403+404
RJH207	pGAD Not11	PCR/restriction cloned	Rebecca Haugen	RJH 401+402
RJH208	pGAD Not2	PCR/restriction cloned	Rebecca Haugen	RJH 364+365
RJH209	pGAD Caf40	PCR/restriction cloned	Rebecca Haugen	RJH 405+406
RJH210	pGAD Not3	PCR/restriction cloned	Rebecca Haugen	RJH 366+367
RJH211	pGAD Not1-N 1-788	invPCR	Rebecca Haugen	RJH 346+347
RJH212	pGAD Not1-N 789-1148	invPCR	Rebecca Haugen	RJH 348+349
RJH213	pGAD Not1-N 789-1064	invPCR	Rebecca Haugen	RJH 348+369
RJH214	pGAD Not1-N 1061-1148	invPCR	Rebecca Haugen	RJH 349+375
RJH215	pGAD Not1-N 789-908	invPCR	Rebecca Haugen	RJH 347+374
RJH216	pGAD Not1-N 908-1064	invPCR	Rebecca Haugen	RJH 349+369
RJH217	pGAD Not1-N 908-1016	invPCR	Rebecca Haugen	RJH 347+449

Archive number	Plasmid name	How cloned	Creator	Primers
RJH218	pGAD Not1-N 908-995	invPCR	Rebecca Haugen	RJH 347+408
RJH219	pGAD Not2 1-399	invPCR	Rebecca Haugen	RJH 347+412
RJH220	pGAD Not2 399-585	invPCR	Rebecca Haugen	RJH 349+411
RJH221	pGAD Not2 399-475	invPCR	Rebecca Haugen	RJH 347+416
RJH222	pGAD Not2 465-585	invPCR	Rebecca Haugen	RJH 349+415
RJH223	pGAD Not3 522-844	invPCR	Rebecca Haugen	RA 221+ RJH 349
RJH224	pGAD Not3 740-844	invPCR	Rebecca Haugen	RJH 457+349
pBA041	pGADT7 RD3		Brian Alzua	
CAW018	pIZ HT-V5H6	Amplified RE site containing HT w/ CW081/082, cut with HindIII and XbaI, and ligated into same cut pIZ vector	Chase Weidmann	CW081/082 for insert AG871/872 in vector
RJH175	pIZ Not1-N FLAG	Amplified from pIZ Not1-N V5; blunt+SpeI digest for insert into pIZ Flag MCS	Rebecca Haugen	RJH 6+171
RJH227	pIZ Not2-Flag	PCR/restriction cloned	Rebecca Haugen	RA 213+272
RJH228	pIZ Not11-V5	PCR/restriction cloned	Rebecca Haugen	RJH 401+402
	pIZ MS2-RD3 V5 H6		Chase Weidmann	
ACG864	pIZ Dm Pumilio V5 His6	PCR Hind3 and Xba1	Aaron Goldstrohm	
CAW003	pIZ DmPumilio R7mut SNVVE to AAVVA	Quickchange PCR	Chase Weidmann	ACG 864 CW015, 016
ACG866	pIZ Dm Pumilio RBD	PCR Hind3 and Xba1	Aaron Goldstrohm	
CAW391	pIZ PumΔRD3 V5H6	PCR amplification and insertion	Chase Weidmann	
ACG544	pIZ GFP		Marzluf Lab	
RMA008	pIZ Dm Pum RD3-RBD	INV PCR/PNK/T4	Rene Arvola	R: CW 055 F: CW 244

Archive number	Plasmid name	How cloned	Creator	Primers
RJH229	pIZ RD3-RBD FF Mut	restriction cloned	Rebecca Haugen	
CAW004	pIZ DmPumRBD R7mut SNVVE to AAVVA	Quickchange PCR	Chase Weidmann	ACG 866 w/ CW015 016
RJH230	pAc 5.4 NLuc 3xPRE (19nt)	subcloned from pNLP w/ restriction sites	Rebecca Haugen	
RJH231	pAc 5.4 NLuc 3xPRE ACA mt (19nt)	subcloned from pNLP w/ restriction sites	Rebecca Haugen	
CAW061	pIZ MS2CP-GFP V5H6	Amplified from EGFP / KpnI and XbaI, cut pIZ MS2CP vector	Chase Weidmann	CW115/116 for insert
RJH232	pAc 5.4 NLuc Chico 3'-UTR	gBlock with restriction sites	Rebecca Haugen	
RJH233	pAc 5.4 NLuc Chico 3'-UTR 2x ACA mt	gBlock with restriction sites	Rebecca Haugen	
RJH234	pAc 5.4 NLuc Nrv1 3'-UTR	gBlock with restriction sites	Rebecca Haugen	
RJH235	pAc 5.4 NLuc Nrv1 3'-UTR 1x ACA mt	gBlock with restriction sites	Rebecca Haugen	
RJH236	pAc 5.4 NLuc Pde11 3'-UTR	gBlock with restriction sites	Rebecca Haugen	
RJH237	pAc 5.4 NLuc Pde11 3'-UTR 2xACA mt	gBlock with restriction sites	Rebecca Haugen	
RJH238	pAc 5.4 NLuc Raf 3'-UTR	gBlock with restriction sites	Rebecca Haugen	
RJH239	pAc 5.4 NLuc Raf 3'-UTR 2xACA mt	gBlock with restriction sites	Rebecca Haugen	
RMA071	pAc-sgRNA-Cas9			
RJH244	pAc sgRNA Cas9 ΔFlag	invPCR	Rebecca Haugen	RJH 234+235
RJH245	pAc sgRNA Cas9 ΔFlag-Blast	amplified Blast from pIZ, restriction cloned w/ XhoI+NheI	Rebecca Haugen	RJH 259+260
RJH246	pAc sgRNA Cas9 dPum sg1	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 526+527
RJH247	pAc sgRNA Cas9 dPum sg2	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 528+529
RJH248	pAc sgRNA Cas9 Not1 sg1	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 212+213

Archive number	Plasmid name	How cloned	Creator	Primers
RJH249	pAc sgRNA Cas9 Not1 sg2	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 210+211
RJH252	pAc sgRNA Cas9 ΔFlag Pum exon13 sg1	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 284+285
RJH253	pAc sgRNA Cas9 ΔFlag Pum exon13 sg2	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 286+287
RJH254	pAc sgRNA Cas9 ΔFlag-Blast Pde11 sg1	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 331+332
RJH255	pAc sgRNA Cas9 ΔFlag-Blast Pde11 sg2	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 333+334
RJH256	pAc sgRNA Cas9 ΔFlag-Blast Raf sg1	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 327+328
RJH257	pAc sgRNA Cas9 ΔFlag-Blast Raf sg2	Oligo cloned with BspQI sites	Rebecca Haugen	RJH 329+330

Table C1: Plasmids used in this thesis.

Appendix D: RT-qPCR primer data

Reactions for RT-qPCR were performed on the Bio-Rad CFX96 instrument using the following cycling parameters: 3 min 98°C, [10 sec 95°C, 30 sec 62/63/64°C, 40 sec 72°C + image] x 39, 60°C-90°C melt curve + image. Individual annealing temps for each primer set are listed below, along with primer efficiency and sequences. Primers are specific for only the intended target unless specified with an asterisk. The non-specific targets for these primers are either not expressed in Dl1 cells (via FlyBase) or have amplicons greater than 800 bp and are not observed on an agarose gel.

Target	FlyBase ID	Fwd Primer	Rev Primer	E value	Ta (°C)	Concen-tration (nM)	Primer sequence F	Primer sequence R	Amplicon length
Act5c	FBgn0038331	RJH 389	RJH 390	1.9413	62	200	TTAGCTCAGCCTC GCCACTT	TCGTCACACATTT TGTAAGCTTTTT	88
Act5c	FBgn0038331	RJH 389	RJH 390	1.9648	64	200	TTAGCTCAGCCTC GCCACTT	TCGTCACACATTT TGTAAGCTTTTT	88
Baboon	FBgn0011300	RJH 476	RJH 477	1.9497	62	200	CTGCAGACAACAA GGACAATGG	GTGCCCACAATGT CCATATGC	178
beta-tubulin	FBgn0003888	RJH 514	RJH 515	1.9956	64	100	AGCCTTCGACCCC ACAAATC	TCCCAGAACTTAG CGCCGAT	97
Ccm3	FBgn0038331	RJH 520	RJH 521	1.9684	62	100	ATCCGGGCTTCTG CTATGAC	TCATTGAGGTCCG CTTCCGT	118
Ches-1-like	FBgn0029504	RJH 502	RJH 503	1.9083	63	200	GTGAGGAGAATCA CAACATCACG	CTGTGATAATGTG TGGTGGCC	136
Chico	FBgn0024248	RJH 484	RJH 485	1.9370	64	200	GTCCAACGTTCAC CATGCC	CGATGACCGCGT GTTTAATCTG	116
CoA7	FBgn0039965	RJH 488	RJH 489	1.9376	63	100	TGGGCGTGGAATA TCGATTCG	CGACTTGTACACC TTGGAGGC	122
CycK	FBgn0025674	RJH 504	RJH 505	1.9290	64	200	GCATTAAATCCCT AGCCAGTGG	AGCATGATAAATG TGGTACCCG	113
E(Pc)	FBgn0000581	RJH 498	RJH 499	1.9447	64	200	CAGCGAGCGATAT GCACG	CCCAAGAGGTTGC ATGTGGATC	132
Fs(1)h	FBgn0004656	RJH 496	RJH 497	1.9359	63	100	GCCAAGAAAGATG AGTCCGCTTC	CTGCTATCACTCG AACTCGACG	113
Larp4B	FBgn0035424	RJH 506	RJH 507	1.9837	64	200	GAGTTGGGCTCTA CGGCAG	TGATTGTACATTG CTCCCGGTG	86
Musashi	FBgn0011666	RJH 480	RJH 481	2.0104	64	200	CCACTGGCCATGC TTCAAAA	GGGTGCATATCTG ACGCTAT	204
Myc	FBgn0262656	RJH 490	RJH 491	1.9276	63	100	AGCGACTGGAAA GCAAAGGA	TGCCTGTCTCGCT GAACATA	90

Not1	FBgn0085436	RJH 512	RJH 513	1.9347	64	200	GGACGTGTGCATGGAACTTGATC	CAGCTGACCTTCCGTGTTTGC	157
Nrg	FBgn0264975	RJH 516	RJH 517	1.9870	62	100	GGCGAGGGATCTGAACACTA	AGTCCATTGTCGGATGGCAG	110
Nrv1	FBgn0015776	RJH 518	RJH 519	1.9671	62	100	TACGGAACCCAAGTGGAAGC	GTGTGGTTGTAGTCCCGGAG	180
Pangolin	FBgn0085432	RJH 474	RJH 475	1.9206	62	100	GCAAAGCACACATCAAATGCG	CTGATTAATCGCAGCCGATTCC	170
Pde11	FBgn0085370	RJH 522	RJH 523	1.9639	62	100	GCCATCAATATCCATACTATCATTGCC	CAAATGTCCTTCACCAGCTCG	112
Pgk	FBgn0250906	RJH 508	RJH 509	2.0177	62	100	GGTGTTGATGCGCGTCGAC	TCTTGTTGCCATCGGGACG	161
Pop2	FBgn0036239	RJH 510	RJH 511	1.9938	62	200	AAGTTTAACCTGAGCGAGGACATG	CAGAGCTCATCAGCAGTTCGG	133
Pop2	FBgn0036239	RJH 510	RJH 511	2.0128	64	200	AAGTTTAACCTGAGCGAGGACATG	CAGAGCTCATCAGCAGTTCGG	133
Pten	FBgn0026379	RJH 486	RJH 487	1.9436	64	200	CAGTTTCCGGCGATGTAAAATTTG	CGGACTCGCAAGGTGAATAG	111
Pum*	FBgn0003165	CW057	CW058	1.9671	62	200	GCCTGATGACCGATGTCTTTGG	CGATTTCCTGCTGCTGCTCC	189
Pum	FBgn0003165	RJH 482	RJH 483	2.0449	64	100	GCCACGTCCTACGTCATCAATCC	GGAATGCCGGGATGACCTGATAC	208
Raf	FBgn0003079	RJH 472	RJH 473	1.9910	62	100	AGACCTCCTTTGCCGCATCC	GATGCGCGGCCCAATTAAAA	195
Rpl32	FBgn0002626	RC 133	RC 134	1.9367	65	100	GCCCAAGGGTATCGACAACA	GCGCTTGTTCGATCCGTAAC	85
Rpl32	FBgn0002626	RC 133	RC 134	1.9263	64	100	GCCCAAGGGTATCGACAACA	GCGCTTGTTCGATCCGTAAC	85
Rpl32	FBgn0002626	RC 133	RC 134	1.9226	63	100	GCCCAAGGGTATCGACAACA	GCGCTTGTTCGATCCGTAAC	85

Rpl32	FBgn0002626	RC 133	RC 134	2.0021	62	100	GCCCAAGGGTATCGACAACA	GCGCTTGTTCGATCCGTAAC	85
Skd	FBgn0003415	RJH 492	RJH 493	1.9571	64	200	CATGGAGAGATTCGTGCTTACC	GAGAAGGACAAGTGCTGGG	114
Spoon	FBgn0263987	RJH 470	RJH 471	1.9553	62	100	CAACATTGAGCCCATCGGTG	ATTATTGGGACCTAGGGATGCG	154
Strica	FBgn0033051	RJH 500	RJH 501	1.9380	63	100	GGAAGTGCCCAGGATGTGAA	CCTTGGTTTGCAACATTCGC	124
Tis11*	FBgn0011837	RC 101	RC 102	1.9224	64	100	GGTCTTGTATCTGGATGTGTTCATCG	TGGGCAACATGAACCTGCAC	92
Tob	FBgn0028397	RJH 494	RJH 495	1.9726	64	200	GCATTGGTGAAAAAGGAGCG	GTTCATGGTGGTATTAACGGCG	152
upSET	FBgn0036398	RJH 478	RJH 479	1.9997	62	200	ATATCCTCCTACATGTCATCGGC	TGCCAACTATGCTATTCACGCC	190
Wac	FBgn0035120	RJH 524	RJH 525	1.9463	62	100	cGACTTCTATTGTGCCAAGAAGAGGG	TTCCGCAGCAGCATTAAATCTCTCC	148

Table D1: qPCR primer data

Appendix E: Fold change data for all tested targets in Chapter 4: Endogenous targets of Pum.

The 23 targets we tested for RNA level changes under Pum depletion are described in Chapter 4. These targets were investigated by two methods. In addition to the RNAi experiments described in Chapter 4, we also used our CRISPR-Cas9 engineered Pum KO lines to evaluate the RNA levels of these targets. These targets were first examined upon generation of our Pum KO lines by RT-qPCR as described. However, after multiple passages of the Pum KO lines, the measured RNA levels did not appear to be affected and we assumed the cells had acclimated to the absence of Pum. In this table we present the measured average fold changes for the first set of experiments in Pum KO lines. Additionally, fold change values are listed for all RNAi experiments. Data is compiled from multiple experiments, including the final three replicates used to generate the figures in Chapter 4.

Gene	Pum KO1 Fold change (avg, n=3)	Pum KO2 Fold change (avg, n=3)	Pum RNAi Fold change (avg)	n= for RNAi
Pum	0.678577398	0.525081489	0.181640346	14
Wac (non-target)	1.559766668	1.048756846	0.911068624	14
Pgk (non-target)			0.88405044	9
β-tubulin (non-target)			0.754258575	9
Fs(1)h	1.5609206	1.318678082	1.041977199	14
Strica	1.2050576	0.670007368		
Larp4B	1.635042	1.396558585	1.059505447	14
Ches-1-L	1.2144014	1.295067598		
Myc	1.530425	1.258629362	1.069917582	4
Coa7	1.6316373	1.648109225	1.158389511	5
Chico	1.316703555	2.122192503	1.213583864	14
Skd	1.080148177	1.283432295		
E(Pc)	1.512280725	1.984834513	1.19411396	13
Msi	1.06661805	1.056450538		
Pten	1.406828321	1.456651015	1.153850923	9
Tob	2.24611848	3.080975635	1.432102831	14
Tis11	1.680963548	1.56625273	1.152147636	14
CycK	1.123255273	0.895703324		
Raf	1.998143104	2.444376559	1.709810008	14
Spoon	1.162615403	1.100030607		
Pan	1.040845778	1.308507525	1.243331993	9
upSET	1.58110159	1.53039345	1.022308817	5
Babo	1.346663396	1.368104056	1.284068113	9
Nrg	1.341054177	1.583092598	1.217768574	11
Nrv1	1.706321357	1.366139905	1.37710822	11
Ccm3	1.882039354	1.753311989	1.22458802	14
Pde11	1.508792832	1.363968501	1.672504459	11

Table E1: Fold change data for targets in Chapter 4.

Appendix F: GO term analyses

Biological Process GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0022008~neurogenesis	155	2.28E-39	3.062539868
GO:0000398~mRNA splicing, via spliceosome	71	1.47E-21	3.507102107
GO:0006364~rRNA processing	28	1.19E-17	6.946147772
GO:0009267~cellular response to starvation	36	7.20E-14	4.230360673
GO:0032543~mitochondrial translation	33	9.00E-13	4.234412743
GO:0000381~regulation of alternative mRNA splicing, via spliceosome	28	4.72E-11	4.281871914
GO:0000462~maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	18	5.17E-11	6.698071066
GO:0006351~transcription, DNA-templated	72	7.55E-10	2.137682255
GO:0006338~chromatin remodeling	21	1.32E-09	4.884010152
GO:0016569~covalent chromatin modification	16	3.98E-09	6.159145808
Molecular Function GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0003723~RNA binding	82	2.84E-22	3.199522565
GO:0003676~nucleic acid binding	93	2.21E-14	2.276354088
GO:0003677~DNA binding	104	5.62E-12	1.98144021
GO:0003729~mRNA binding	48	7.39E-12	2.968149303
GO:0030515~snoRNA binding	12	4.82E-09	8.459225513
GO:0005515~protein binding	106	1.05E-08	1.732376168
GO:0000166~nucleotide binding	48	1.75E-08	2.394120428
GO:0003682~chromatin binding	35	1.95E-08	2.868923381
GO:0042393~histone binding	14	1.41E-07	5.69370948
GO:0003899~DNA-directed RNA polymerase activity	14	4.25E-07	5.287015945

Table F1: Gene Ontology analysis for upregulated genes in Pum KO lines illustrated in Figure 5.8. The top ten significantly enriched terms ranked by p-value for each category are listed. Count is the number of genes that fall into the listed category.

Biological Process Go term	Count	P Value	Fold Enrichment
GO:0007476~imaginal disc-derived wing morphogenesis	29	4.64E-11	4.493919
GO:0006351~transcription, DNA-templated	34	2.13E-08	3.040731
GO:0022008~neurogenesis	43	3.05E-08	2.559216
GO:0045892~negative regulation of transcription, DNA-templated	18	8.29E-08	5.044037
GO:0007411~axon guidance	20	2.37E-06	3.635342
GO:0030097~hemopoiesis	8	3.07E-06	11.69632
GO:0006355~regulation of transcription, DNA-templated	30	8.16E-06	2.515729
GO:0000122~negative regulation of transcription from RNA polymerase II promoter	17	8.48E-06	3.836627
GO:0006338~chromatin remodeling	10	9.58E-06	7.005607
GO:0006357~regulation of transcription from RNA polymerase II promoter	18	1.46E-05	3.498754
Molecular Function GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0005515~protein binding	44	6.40E-07	2.246866
GO:0003677~DNA binding	38	3.96E-06	2.26214
GO:0003682~chromatin binding	14	1.33E-04	3.585644
GO:0003676~nucleic acid binding	27	5.87E-04	2.064947
GO:0008375~acetylglucosaminyltransferase activity	4	0.002782	13.21566
GO:0003723~RNA binding	18	0.003395	2.194482
GO:0001047~core promoter binding	4	0.003741	12.01423
GO:0008134~transcription factor binding	9	0.004608	3.417843
GO:0042393~histone binding	5	0.00717	6.353682
GO:0003729~mRNA binding	12	0.01437	2.318537

Table F2: Gene Ontology analysis for PRE-containing upregulated genes in Pum KO lines. The top ten significantly enriched terms ranked by p-value for each category are listed. Count is the number of genes that fall into the listed category.

Biological Process GO term	Count	P Value	Fold Enrichment
GO:0006897~endocytosis	21	4.48E-06	3.130088
GO:0007274~neuromuscular synaptic transmission	19	1.26E-04	2.703258
GO:0008039~synaptic target recognition	12	3.99E-04	3.414642
GO:0007616~long-term memory	19	4.13E-04	2.477987
GO:0007165~signal transduction	34	4.40E-04	1.878053
GO:0045886~negative regulation of synaptic growth at neuromuscular junction	14	5.89E-04	2.921416
GO:0055085~transmembrane transport	50	6.83E-04	1.619011
GO:0000045~autophagosome assembly	10	7.12E-04	3.756106
GO:0006749~glutathione metabolic process	14	9.28E-04	2.7971
GO:0035071~salivary gland cell autophagic cell death	19	0.001143	2.287372
Molecular Function GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0003779~actin binding	29	3.67E-05	2.282899
GO:0016491~oxidoreductase activity	41	9.90E-05	1.872981
GO:0005215~transporter activity	25	9.67E-04	2.022182
GO:0003997~acyl-CoA oxidase activity	5	0.002037	7.347262
GO:0004872~receptor activity	18	0.002737	2.173984
GO:0019901~protein kinase binding	12	0.003889	2.645014
GO:0005509~calcium ion binding	38	0.004041	1.595405
GO:0017137~Rab GTPase binding	16	0.006038	2.137385
GO:0005089~Rho guanyl-nucleotide exchange factor activity	9	0.006373	3.05194
GO:0003924~GTPase activity	26	0.007264	1.723568

Table F3: Gene Ontology analysis for downregulated genes in Pum KO lines illustrated in Figure 5.8. The top ten significantly enriched terms ranked by p-value for each category are listed. Count is the number of genes that fall into the listed category.

Biological Process GO term	Count	P Value	Fold Enrichment
GO:0008039~synaptic target recognition	9	8.62E-06	8.261458
GO:0007274~neuromuscular synaptic transmission	11	5.38E-05	5.048669
GO:0007165~signal transduction	17	1.52E-04	3.029201
GO:0007411~axon guidance	16	0.00117	2.61985
GO:0048675~axon extension	6	0.002327	6.267313
GO:0007616~long-term memory	9	0.002432	3.786501
GO:0050770~regulation of axonogenesis	5	0.00245	8.414448
GO:0007523~larval visceral muscle development	4	0.002564	13.46312
GO:0098655~cation transmembrane transport	4	0.002564	13.46312
GO:0055085~transmembrane transport	20	0.003372	2.089104
Molecular Function GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0003779~actin binding	15	2.48E-05	3.922375
GO:0005509~calcium ion binding	20	9.56E-05	2.789244
GO:0019901~protein kinase binding	8	3.51E-04	5.857413
GO:0005089~Rho guanyl-nucleotide exchange factor activity	6	0.001615	6.758554
GO:0042626~ATPase activity, coupled to transmembrane movement of substances	7	0.003124	4.767662
GO:0004889~acetylcholine-activated cation-selective channel activity	4	0.003922	11.71483
GO:0005515~protein binding	34	0.012396	1.539042
GO:0051015~actin filament binding	5	0.014143	5.229833
GO:0050839~cell adhesion molecule binding	5	0.015977	5.049494
GO:0005096~GTPase activator activity	8	0.028412	2.693064

Table F4: Gene Ontology analysis for PRE-containing downregulated genes in Pum KO lines. The top ten significantly enriched terms ranked by p-value for each category are listed. Count is the number of genes that fall into the listed category.

Biological Process GO term	Count	P Value	Fold Enrichment
GO:0007476~imaginal disc-derived wing morphogenesis	29	1.08E-10	4.334842
GO:0006351~transcription, DNA-templated	36	4.25E-09	3.10563
GO:0022008~neurogenesis	44	3.01E-08	2.526034
GO:0045892~negative regulation of transcription, DNA-templated	18	1.40E-07	4.865487
GO:0007411~axon guidance	21	1.00E-06	3.68199
GO:0006357~regulation of transcription from RNA polymerase II promoter	20	1.47E-06	3.749893
GO:0030097~hemopoiesis	8	3.91E-06	11.28229
GO:0006355~regulation of transcription, DNA-templated	31	5.94E-06	2.507566
GO:0006338~chromatin remodeling	10	1.28E-05	6.75762
GO:0000122~negative regulation of transcription from RNA polymerase II promoter	17	1.33E-05	3.700818
Molecular Function GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0005515~protein binding	47	9.67E-08	2.317585
GO:0003677~DNA binding	39	3.59E-06	2.241887
GO:0003682~chromatin binding	14	1.89E-04	3.462426
GO:0003676~nucleic acid binding	27	9.92E-04	1.993986
GO:0008375~acetylglucosaminyltransferase activity	4	0.003075	12.76151
GO:0001047~core promoter binding	4	0.004131	11.60137
GO:0003723~RNA binding	18	0.004889	2.11907
GO:0008134~transcription factor binding	9	0.005678	3.300391
GO:0042393~histone binding	5	0.008102	6.135342
GO:0003729~mRNA binding	12	0.018201	2.238862

Table F5: Gene Ontology analysis for the 417 Top Pum targets illustrated in Figure 5.12. The top ten significantly enriched terms ranked by p-value for each category are listed. Count is the number of genes that fall into the listed category.

Biological Process GO term	Count	P Value	Fold Enrichment
GO:0007298~border follicle cell migration	41	5.13E-15	4.146381
GO:0006886~intracellular protein transport	31	1.13E-10	3.874471
GO:0016567~protein ubiquitination	36	5.00E-10	3.266678
GO:0043087~regulation of GTPase activity	20	1.24E-09	5.299277
GO:0032456~endocytic recycling	13	2.23E-09	8.611325
GO:0016192~vesicle-mediated transport	30	1.27E-08	3.312048
GO:0006367~transcription initiation from RNA polymerase II promoter	21	2.28E-08	4.347063
GO:0045746~negative regulation of Notch signaling pathway	18	3.19E-08	4.968072
GO:0030036~actin cytoskeleton organization	17	7.62E-08	5.004873
GO:0016322~neuron remodeling	17	2.14E-07	4.692068
Molecular Function GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0005515~protein binding	114	5.26E-17	2.231679
GO:0005096~GTPase activator activity	28	2.50E-10	4.076336
GO:0004842~ubiquitin-protein transferase activity	39	1.18E-08	2.759578
GO:0017137~Rab GTPase binding	22	1.52E-08	4.221919
GO:0051015~actin filament binding	12	4.19E-06	5.428182
GO:0001104~RNA polymerase II transcription cofactor activity	13	5.71E-06	4.84279
GO:0003779~actin binding	23	5.12E-05	2.601004
GO:0003713~transcription coactivator activity	15	9.97E-05	3.333094
GO:0004702~receptor signaling protein serine/threonine kinase activity	10	1.03E-04	4.871445
GO:0004672~protein kinase activity	27	1.44E-04	2.235134

Table F6: Gene Ontology analysis for upregulated genes in Not1 and Pop2 RNAi conditions illustrated in Figure 5.10. The top ten significantly enriched terms ranked by p-value for each category are listed. Count is the number of genes that fall into the listed category.

Biological Process GO term	Count	P Value	Fold Enrichment
GO:0002181~cytoplasmic translation	45	5.72E-24	6.049144
GO:0006457~protein folding	24	3.32E-06	2.987232
GO:0005975~carbohydrate metabolic process	23	8.41E-06	2.916778
GO:0015992~proton transport	11	8.60E-06	5.68723
GO:0006096~glycolytic process	11	1.28E-05	5.476592
GO:0045454~cell redox homeostasis	15	1.72E-05	3.877657
GO:0015986~ATP synthesis coupled proton transport	10	3.12E-05	5.601059
GO:0006486~protein glycosylation	14	1.48E-04	3.421738
GO:0006364~rRNA processing	12	3.47E-04	3.584678
GO:0006488~dolichol-linked oligosaccharide biosynthetic process	5	4.02E-04	11.20212
Molecular Function GO term	**Count**	**P Value**	**Fold Enrichment**
GO:0003735~structural constituent of ribosome	58	1.00E-08	2.223979
GO:0046933~proton-transporting ATP synthase activity, rotational mechanism	10	3.81E-06	6.795491
GO:0003723~RNA binding	42	7.76E-05	1.895716
GO:0019843~rRNA binding	10	1.88E-04	4.530327
GO:0003729~mRNA binding	29	2.80E-04	2.074413
GO:0005524~ATP binding	88	4.16E-04	1.433297
GO:0000030~mannosyltransferase activity	5	0.001263	8.73706
GO:0046961~proton-transporting ATPase activity, rotational mechanism	11	0.00135	3.281725
GO:0003756~protein disulfide isomerase activity	7	0.002293	4.756844
GO:0051287~NAD binding	10	0.002375	3.305915

Table F7: Gene Ontology analysis for downregulated genes in Not1 and Pop2 RNAi conditions illustrated in Figure 5.10. The top ten significantly enriched terms ranked by p-value for each category are listed. Count is the number of genes that fall into the listed category.

Appendix G: CRISPR data

Gene	HR template
Pum	ACAGCAGCTTGGGTCCCATTGGACCCCCGACCAACGGCAACGTTGTGct ggaggtgctgttccagggccccgaacaaaaactcatctcagaagaggatctgTAAAGGAAATAA CAAATTAAGCCAAGCAGTCAAAGGAAACTTCTTTCTCGAATCGCAGTATA GTTTTTAGAAGCTGTAGAGCTTAACATAAACAACAAG
Not1	CGTCGATGGCGAGGGCCAGGAGGTAGCCACCATCAACcttgaggatctgtacttt cagagccttgactacaaagaccatgacggtgattataaagatcatgacatcgattacaaggatgacgatg acaagTGAATGGATCCACGTCCGACGCTTCTGCCTGGAGTTCTGCGCGA GACCCAGACGCAGGCAGTAGCTGCCT
Pde11	TGTATGAATTCGAATGTCCTGATAATGCGAATTAAGTATTTATTTTTCCAT TACTCCTTCTTCGTTGCAGGCggtaagcctatccctaaccctctcctcggtctcgattctacg CTGCATCTGCGGAGCATCGAAGAGCCAGCGACGACGCCGCTCCAGTTC CAGCCCACTGGCAGGTGAACACGGAGCAGGGCGGCAC
Raf	TTCGGGGTCATGGTCACAGCGCATAGTATATAGGATAAAGCAACACCAT GggtaagcctatccctaaccctctcctcggtctcgattctacgTCCAGCGAGTCCAGCACCGA AGGCGACAGCGATCTATACGATCCTTTGGCCGAGGAGCTGCACAACGT CCAGCTCGTCAAACATGTGACCCGCGAGAATATTGATGCC

Table G1: HR template sequences. HR templates as ordered from IDT for CRISPR-Cas 9 targeting. Lower case indicates sequence of the intended tag, while bases in all caps are homology arms.

Primer name	Sequence
RJH 210 Not1 SG2 Fwd	ttcGCGTCCGACGCTTCTGCCTGG
RJH 211 Not1 SG2 Rev	aacCCAGGCAGAAGCGTCGGACGC
RJH 212 Not1 Sg1 Fwd	ttcgCTCCAGGCAGAAGCGTCGGA
RJH 213 Not1 SG1 Rev	aacTCCGACGCTTCTGCCTGGAGc
RJH 284 Pum exon 13 sg1 Fwd	ttcgAGGAAATAACAAATTAAGCC
RJH 285 Pum exon 13 sg1 Rev	aacGGCTTAATTTGTTATTTCCTc
RJH 286 Pum exon 13 sg2 Fwd	ttcgAACTATACTGCGATTCGAGA
RJH 287 Pum exon 13 sg2 Rev	aacTCTCGAATCGCAGTATAGTTc
RJH 327 Raf sg1 Fwd	ttcgTAGATCGCTGTCGCCTTCGG
RJH 328 Raf sg1 Rev	aacCCGAAGGCGACAGCGATCTAc
RJH 329 Raf sg2 Fwd	ttcGGAGCTGCACAACGTCCAGT
RJH 330 Raf sg2 Rev	aacACTGGACGTTGTGCAGCTCC
RJH 331 Pde11 sg1 Fwd	ttcGCAGATGCAGGCCTGCAACG
RJH 332 Pde11 sg1 Rev	aacCGTTGCAGGCCTGCATCTGC
RJH 333 Pde11 sg2 Fwd	ttcgCCTGCATCTGCGGAGCATCG
RJH 334 Pde11 sg2 Rev	aacCGATGCTCCGCAGATGCAGGc
RJH 526 Pum exon 9 crispr 1 Fwd	ttcgCACATCGGCAGCGAGTGCTG
RJH 527 Pum exon 9 crispr 1 Rev	aacCAGCACTCGCTGCCGATGTGc
RJH 528 Pum exon 9 crispr 2 Fwd	ttcGCAACAACAGATGCACATGG
RJH 529 Pum exon 9 crispr 2 Rev	aacCCATGTGCATCTGTTGTTGC

Table G2: Primer sequences for sgRNA used in CRISPR-Cas9 editing. The first “ttc” or “aac” are overhangs for BspQI cloning. An additional lowercase “g” or “c” is appended to some sequences due to the requirement of the U6 promoter, which requires the first base of the transcribed sequence to be guanine.